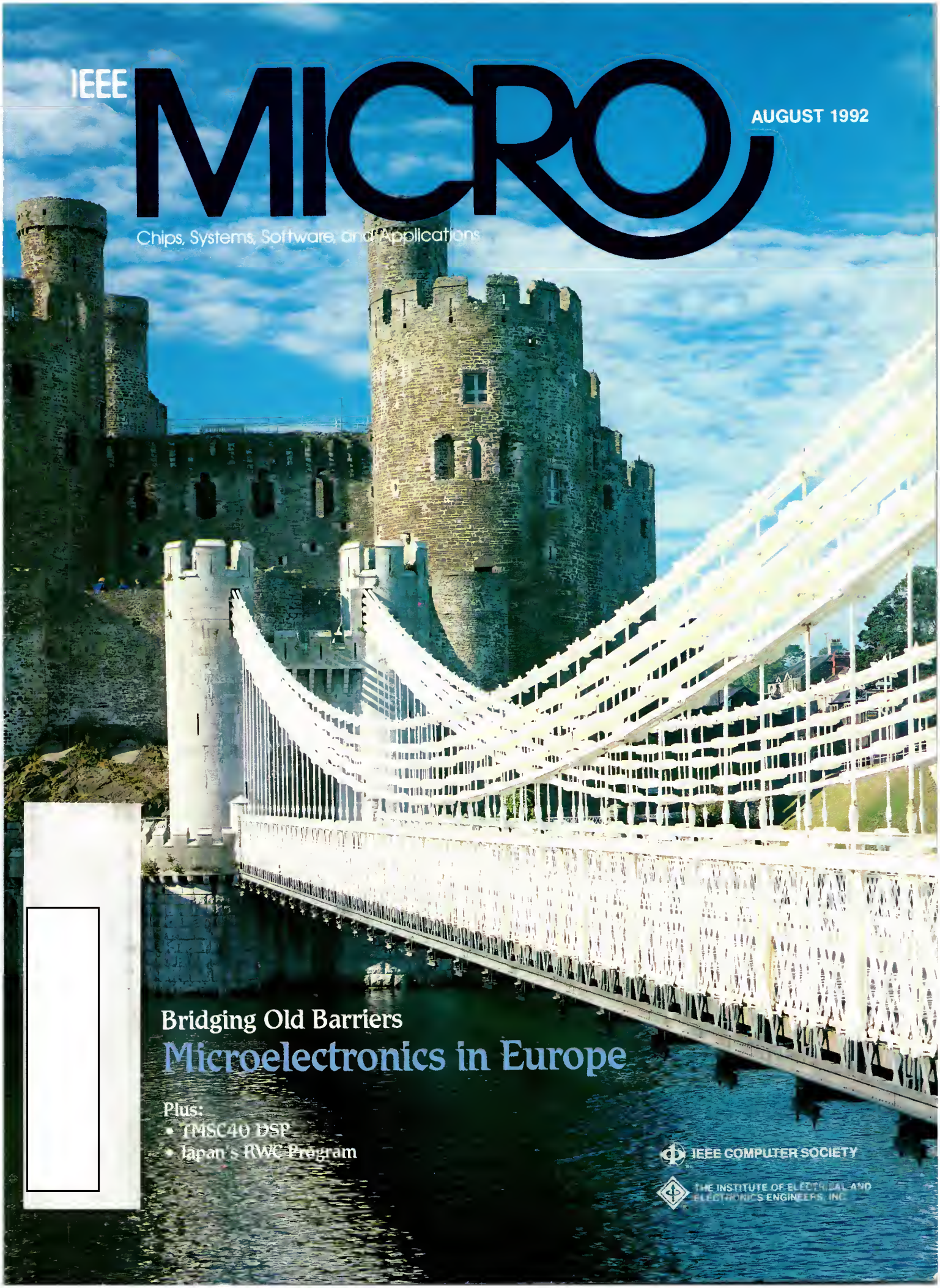
IEEE MICRO

AUGUST 1992

Chips, Systems, Software, and Applications



Bridging Old Barriers

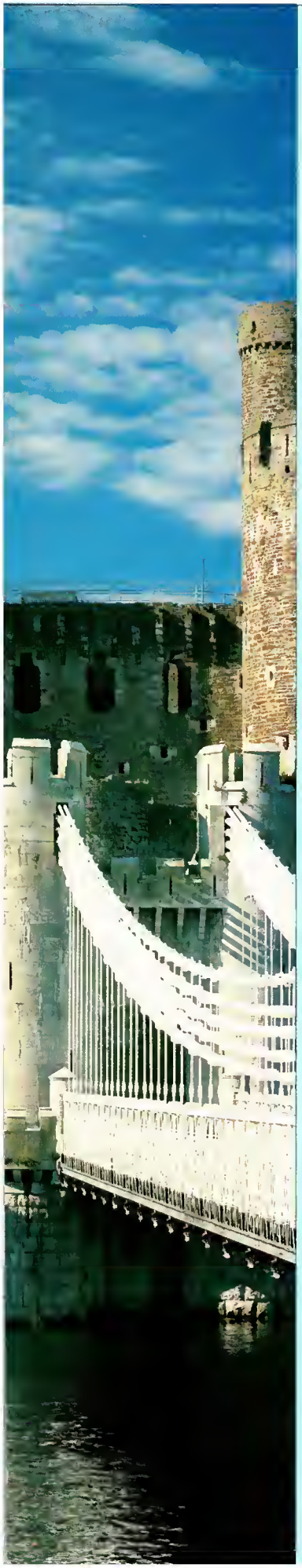
Microelectronics in Europe

Plus:

- TMSC40 DSP
- Japan's RWC Program

IEEE COMPUTER SOCIETY

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC



IEEE MICRO

Published by the IEEE Computer Society

August 1992

FEATURES

Cover Image: FPG International

Cover design: Chris Martin, Design and Direction

Circulation: *IEEE Micro* (ISSN 0272-1732) is published bimonthly by the IEEE Computer Society, PO Box 3014, Los Alamitos, CA 90720-1264; IEEE Computer Society Headquarters, 1730 Massachusetts Ave., NW, Washington, DC 20036-1903; IEEE Headquarters, 345 East 47th St., New York, NY 10017. Annual subscription: $23 in addition to IEEE Computer Society or any other IEEE society member dues; $39 for members of other technical organizations. This journal is also available in microfiche form.
Postmaster: Send address changes and undelivered copies to *IEEE Micro*, PO Box 3014, Los Alamitos, CA 90720-1264. Second-class postage is paid at New York, NY, and at additional mailing offices. Canadian GST#125634188.

DEPARTMENTS

IEEE Computer Society
PO Box 3014, Los Alamitos, CA 90720-1264
(714) 821–8380

Editorial: Unless otherwise stated, bylined articles and descriptions of products and services reflect the author's or firm's opinion; inclusion in this publication does not necessarily constitute endorsement by the IEEE or the IEEE Computer Society. Send editorial correspondence to *IEEE Micro*, PO Box 3014, Los Alamitos, CA 90720-1264. All submissions are subject to editing for style, clarity, and space considerations.

* Submit six copies of all articles and special-issue proposals to Dante Del Corso, Dipartimento di Elettronica, Politecnico di Torino, C.so Duca degli Abruzzi, 24, 10129 Torino, Italy; phone +39-11-556-4044; Compmail: d.delcorso; Internet: delcorso@polito.it; Bitnet: delcorso@itopoli
or
** Ashis Khan, Mips Computer Systems, Inc., 950 DeGuigne Drive, Sunnyvale, CA 94086; Internet: ashis@mips.com.

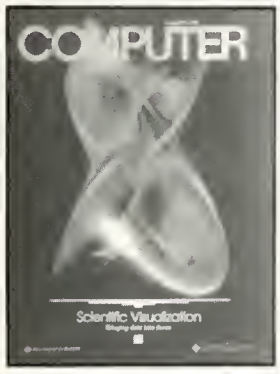
Author guidelines for article submission are available from the West Coast office (address at left).

IEEE COMPUTER SOCIETY

1992 Membership/Subscription Application

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

BENEFITS



Receive *Computer* automatically with membership. Written and refereed by experts, it features articles on the latest developments in computer technology, applications, and research, as well as survey and tutorial articles on topics covering the entire computer field.

Computer's popular departments include New Products, Product Reviews, Book Reviews, Standards, Open Channel (a reader forum), and Computer Society News.

Other membership benefits include the following:

- Discounts of up to 50% on **Computer Society Press books and videos.** Over 700 titles covering a broad spectrum of computer science topics including networking, communications, advanced systems, image processing, security, artificial intelligence, and visualization
- Discounts on registration to over 100 conferences, workshops, and tutorials each year
- Low subscription rates on special-interest periodicals
- Opportunity to participate in any of our 32 technical committees—networks of professionals with common interests in computer hardware, software, and applications
- Opportunity to participate in the development of more than 200 standards projects sponsored by the society
- Low member rates on COMPMAIL, the electronic mail service especially designed for computer professionals

Schedule of Fees

To join: see item 1, 2, or 3.
To subscribe: see item 4.

Membership dues and periodical subscriptions are annualized to, and expire on, December 31. Pay full- or half-year rate depending on date of receipt by the Computer Society as indicated below.

PRICES EXPIRE 12/31/92

	Half Year Mar 1-Aug 31	Full Year Sept 1-Feb 28
1 I don't belong to the IEEE and I want to join just the Computer Society	☐ $27	☐ $ 54
2 I want to join both the Computer Society and the IEEE*		
I reside in Region 1-6 (United States)	☐ $58.50	☐ $117
I reside in Region 7 (Canada)	☐ $53.50	☐ $107
I reside in Region 8 (Europe, Africa, or the Middle East)	☐ $53.00	☐ $106
I reside in Region 9 (Latin America)	☐ $49.50	☐ $ 99
I reside in Region 10 (Asia and Pacific)	☐ $48.50	☐ $ 97
3 I already belong to the IEEE and I want to join the Computer Society	☐ $11.00	☐ $22
IEEE Member Number _____		

*ACM members who join both IEEE and the Computer Society may deduct $5 off the full-year rate; $2.50 off the half-year rate.

4 SPECIAL-INTEREST PERIODICALS for new or current members

	issues per year		
IEEE Computer Graphics and Applications	6	☐ $12.50	☐ $ 25
IEEE Design and Test	4	☐ $11.00	☐ $ 22
IEEE Expert	6	☐ $10.00	☐ $ 20
IEEE Micro	6	☐ $11.50	☐ $ 23
IEEE Software	6	☐ $12.50	☐ $ 25
Transactions on:			
Computers	12	☐ $12.00	☐ $ 24
Knowledge and Data Engineering	6	☐ $ 9.50	☐ $ 19
Parallel and Distributed Systems	6	☐ $ 9.50	☐ $ 19
Pattern Analysis and Machine Intelligence	12	☐ $12.00	☐ $ 24
Software Engineering	12	☐ $11.00	☐ $ 22
IEEE Annals of the History of Computing	4	☐ $ 8.00	☐ $ 16

☐ Visa ☐ MasterCard ☐ American Express

Charge Card Number (Minimum Charge $10)

Mo. Yr.

Exp. Date

Checks drawn in local currency on a bank in the country of origin accepted from members in Australia, Austria, Belgium, Canada, Denmark, France, Germany, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, UK, USA, and Yugoslavia.

Periodicals total $_____

CA, DC residents add applicable sales tax $_____

Membership fees $_____

Total $_____

CANADIAN residents add 7% GST; BELGIAN residents add 6% VAT $_____

Total enclosed $_____

I hereby make application for Computer Society membership and, if elected, will be governed by IEEE's and the society's constitutions, bylaws, and statements of policies and procedures.

Full signature _____ Date _____

☐ Male ☐ Female

MAILING ADDRESS

First name _____ Middle initial(s) _____ Last name _____ Date of birth _____

Street address _____ City _____ State/Country _____ Zip _____

EDUCATION (highest level completed) Course _____ Degrees received _____ Date _____

Name of educational institution _____

OCCUPATION Title or Position _____

Firm name _____

Firm address _____

City _____ State/Country _____ Zip _____

ENDORSER (an IEEE member, Senior Member, or Fellow)

Endorser's signature _____

Name (print in full) _____ IEEE Member No. _____

Street address _____

City _____ State/Country _____ Zip _____

Mail or fax to appropriate Computer Society office:

EUROPE: 13, Avenue de l'Aquilon, B-1200, Brussels, BELGIUM. Fax: 32-2-770-8505
PACIFIC RIM: Ooshima Building, 2-19-1 Minami-Aoyama, Minato-ku, Tokyo 107, JAPAN. Fax: 81-3-3408-3553
ALL OTHERS: 10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720–1264 USA. Fax: 714-821-4010

Current developments in US computer and electronics-industry tie-in law

Richard H. Stern

Oblon, Spivak, McClelland, Maier & Neustadt, P.C.

1755 Jefferson Davis Highway, Suite 400

Arlington, VA 22202

The last 10 or 15 years have probably been an all-time low period for federal antitrust enforcement. The Executive Branch has been remarkably hostile to antitrust enforcement and an apparent majority of judicial appointees during this period have shared that sentiment. Consequently, the Department of Justice stopped more than token enforcement of the antitrust laws. It has also frequently urged the federal courts (which in the main needed little persuasion) to give similar shrift to antitrust cases brought by private parties. This period has marked a withering away of the antitrust laws to the point that they have come to resemble a vermiform appendix of the law, or a quaint curiosity from an earlier time. A recent decision of the US Supreme Court, however, suggests that perhaps the antitrust laws have now passed through the bottom of their parabola, and their dried bones may yet live again.

Virtual Maintenance v. Prime Computer

The Supreme Court's decision can best be understood by placing it in the context of what has usually been going on in the field. Typical of recent antitrust decisions is that of the US Court of Appeals for the Sixth Circuit, in February of this year, in *Virtual Maintenance, Inc.* v. *Prime Computer, Inc.*, 957 F.2d 1318 (6th Cir. 1992) This decision overturned a $25 million-plus jury verdict in favor of an independent service provider and against a computer manufacturer.

Prime manufactures and markets the Prime Series 50 minicomputer system. Ford Motor Company developed PDGS, a CAD/CAM program for Ford equipment suppliers. Ford wrote PDGS to run on Prime 50 minicomputers and licensed Prime as the exclusive distributor of the program. Prime distributes updates, modifications, and revisions ("software support") for PDGS software, which it offers to companies supplying products to Ford, under two plans. In Plan 1 Prime offers a package of PDGS software support and hardware maintenance, for $16,000 per year per Prime 50 system. In Plan 2 Prime offers software support without accompanying hardware maintenance, for $80,000 to $160,000 per year per Prime 50 system. Approximately 350 to 400 out of the installed base of 23,000 Prime Series 50 systems can use PDGS and are thus potential customers for software support and maintenance services. Ford has standardized its operations on the PDGS CAD/CAM program and requires all of its vendors to use the most recent version of PDGS as a condition of doing business with Ford.

Virtual Maintenance sought to enter the hardware maintenance market for Prime 50 systems, after successfully penetrating the markets of two other manufacturers. But VM found it difficult to persuade customers to switch to Prime's Plan 2, under which customers pay five to 10 times as much for software alone as they do under Prime's Plan 1 for a software/maintenance package. None of the customers could do without PDGS, for that would mean giving up Ford's business. One Ford supplier testified at trial on this point. He was asked what the cost would be to switch from Prime computers to DEC or Hewlett-Packard systems:

> A: We can't do that unless we want to say goodbye, unless we want to break our contracts with Ford.
>
> Q: Let's say for some reason your company decided it wanted to say goodbye to those tens of millions of dollars with Ford and they just don't want Prime any

0272-1732/92/0800-0003$03.00 © 1992 IEEE

more. What would be the costs to switch all that to another CAD/CAM system?

A: Well it would be the $2 million investment all over again. But I don't think we would ever do that because it would be business insanity.

VM sued Prime for engaging in an illegal tie-in between software updates and hardware maintenance. The jury found in favor of VM, and the district court denied Prime's motion for a judgment notwithstanding the jury verdict. Prime then appealed to the Sixth Circuit, which reversed.

The court ruled that a tie-in existed but held that the tie-in was not illegal because Prime lacked sufficient market power to cause an appreciable effect on competition. VM argued that Prime must be considered to have market power, because many customers yielded to its tie-in. The court disagreed. It said that VM must show not just that Prime's tie-in program was successful but that Prime is able to force customers to buy the software/maintenance package *because of Prime's market power over the software*. That meant that VM must show that Prime's sales of the tying product (the PDGS software updates) accounted for a substantial fraction of sales in the "relevant market."

The district court gave the jury two choices for the relevant market: software support for all CAD/CAM and software support for software necessary to do business with Ford. The Sixth Circuit considered the first possible relevant market "legally adequate," but it found that Prime had only an 11 percent share of that relevant market. It ruled, as a matter of law, that 11 percent "is insufficient to confer market power." That some customers want PDGS did not give Prime market power, because PDGS software amounts to less than 3 percent of all CAD/CAM software.

To be sure, there was an alternative theory of relevant market, on which the jury could seek to base its verdict. But a new trial was unnecessary, the court said, because the relevant market based on software required to do business with Ford was legally untenable as a relevant market. That is, a relevant market cannot be based on one brand of product. VM argued that Ford's insistence on PDGS defined a relevant market for vendors who wanted to sell to Ford. But the court said this improperly defined a market in terms of the demand side of the market. That approach failed to consider the supply side of the market, which was an obligatory consideration.

The court considered the supply side of the market and concluded that the only legally tenable relevant market for the tying product (PDGS updates) was all CAD/CAM software. Even though Prime is the only supplier of PDGS updates, it lacks any market power over price in that broadly defined market. If Prime charged exorbitant prices, the court theorized, "its customers would simply switch" to some other kind of software. Hence, the court reasoned, Prime cannot charge exorbitant prices and thus lacks true market power.

VM countered that Ford and its vendors cannot just switch to an alternative $2 million minicomputer system, because they are "locked in" once they have invested in Prime 50 units. "Ford presumably had many companies (such as DEC and HP) to choose from prior to purchasing Prime's equipment," the court said, and therefore Prime's power to charge high prices is constrained by the existence of actual or potential competitors. "There exists at least a potential for reasonable interchangeability of supply."

On the basis of these theories and speculations about how the unseen hand of the market is operating, the Sixth Circuit concluded that the jury verdict "ignores economic reality." Finally, the court held, Prime's foreclosure of 400 minicomputer systems in the worldwide market for computer hardware maintenance "is insignificant as a matter of law." (This insignificant business amounted to a pittance of $500 million per year.) Accordingly, the Sixth Circuit reversed the jury verdict and instructed the district court to enter a judgment in favor of Prime.

Prime's foreclosure of 400 minicomputer systems amounted to a pittance of $500 million per year.

Eastman Kodak *v.* Image Technical

In June 1992, the Supreme Court reached a contrary result on similar facts, in *Eastman Kodak Co.* v. *Image Technical Services, Inc.* Image was one of 18 independent service organizations (ISOs) that sued Kodak for tying repair parts to maintenance services for Kodak copying and microfilm machines. The district court granted summary judgment for Kodak on basically the same theory as that stated by the Sixth Circuit in the *Virtual Maintenance* case. The ISOs appealed to the Ninth Circuit, which reversed, and the Supreme Court reviewed the case "because of the importance of the issue."

Kodak provides service and parts for its machines. It manufactures some of the parts and had independent original equipment manufacturers (OEMs) manufacture the rest. The ISOs began to chip away at the market for service of this equipment in the 1980s, although Kodak still accounts for 80 per-

cent or more of the service for Kodak machines. The ISOs charge less than Kodak, and some customers consider their service better. Some customers purchase their own parts and have ISOs install them. Some customers have ISOs supply parts and service. ISOs keep an inventory of parts which they purchase from Kodak, the OEMs, and other sources.

In the mid-1980s Kodak instituted a policy that it would sell repair and replacement parts for Kodak equipment only to owners of Kodak equipment who 1) used Kodak service or 2) performed their own repairs. Kodak would not sell to ISOs or customers using ISOs. In addition, Kodak agreed with OEMs that they would not sell parts for Kodak equipment to anyone but Kodak itself. Kodak succeeded in driving many of the ISOs out of business, which led to this antitrust suit.

The ISOs claimed that Kodak unlawfully tied the sale of service for Kodak machines to the sale of Kodak parts, in violation of section 1 of the Sherman Act. They also claimed that Kodak had illegally monopolized and attempted to monopolize the sale of service for Kodak machines, in violation of section 2 of the Sherman Act. The district court threw the whole case out on Kodak's motion for summary judgment.

The Supreme Court reversed, holding that Kodak had not met the burden of proof on a motion for summary judgment. A jury or other fact finder would have to decide the facts. The Court rejected Kodak's various arguments as to why it was entitled to a judgment in its favor, as a matter of law, without need for a trial.

Market power was the main issue. The ISOs showed that some parts were available only through Kodak, and that Kodak had succeeded in getting OEMs to agree not to sell Kodak parts. Kodak used its power over the market for Kodak parts to exclude ISO competition, boost service prices, and force customers to obtain their service from Kodak instead of ISOs. The ISOs produced evidence that some customers switched to Kodak from ISOs even though 1) they preferred the ISOs' service, 2) Kodak charged higher prices for service, and 3) Kodak's service was of poorer quality than that of the ISOs.

Kodak argued that it lacked the kind of market power required to find a violation of the Sherman Act, because competition exists in the equipment market, where Xerox and IBM compete vigorously against Kodak. Kodak said that if it raised prices above a competitive level, customers would switch to other equipment and it would lose in equipment sales whatever it gained in the parts market. Kodak asked the Court to adopt a rule of law that "equipment competition precludes any finding of monopoly power in derivative aftermarkets." Unless the Supreme Court accepted this rule, Kodak argued, the antitrust laws would deter Kodak from engaging in procompetitive behavior, a factor considered important in the 1986 *Matsushita* case (*Matsushita Elec. Indus. Co.* v. *Zenith Radio Corp.*, 475 U.S. 574 (1986).

Market power was the main issue.

In the *Matsushita* case, the Court threw out a claim that 20 Japanese electronics companies had been engaged in a 20-year conspiracy to sell below cost to gain a monopoly and ultimately raise prices. The Supreme Court said that it would make no sense for the defendants to sell below cost for 20 years in the hope of ultimately making a profit because of having done so. Therefore, the Supreme Court concluded, the plaintiffs' legal theory was unbelievable, and no reasonable jury would be able to rule in their favor.

But the *Matsushita* rule did not apply here, the Court said, because it was not a rule that the defendant is supposed to win if it "enunciates any economic theory supporting its behavior, regardless of its accuracy in reflecting the actual market." Here, Kodak—supported by the Department of Justice as amicus curiae (or independent "friend of the court")—seeks to have the plaintiff ISOs' case thrown out on the theory that Kodak would not have done what the plaintiffs' evidence purports to show that they did. They reasoned that Kodak could not have made money doing it.

Kodak said it would necessarily have lost more on equipment sales decreases than it would have gained on parts gouges. But that was just Kodak's unsupported assertion, which is not necessarily true, the Court said, as a matter of economic theory. Kodak's theory also did not coincide with the evidence that Kodak did charge high prices and did not lose equipment sales. Further, Kodak's theory did not take into account the high cost of switching to a competitive brand, once a customer had purchased Kodak equipment.

As for Kodak's warning that any ruling against it would deter it from engaging in procompetitive behavior, the Court said that the present case was a far cry from *Matsushita*, where that argument was accepted. In *Matsushita*, the defendants were sued for price cutting, which the Court considered was ordinarily procompetitive and desirable behavior that should not be deterred or "chilled." Here, Kodak is raising prices, which is not ordinarily procompetitive, even though the Department of Justice has apparently now taken to arguing that it is. Basically, Kodak's arguments and those of the Department of Justice in Kodak's support are appeals to embrace theory, as handed down from the heights of the University of Chicago's Department of Economics, and to disregard any facts that contradict those theories.

continued on p. 95

Books and more books

Richard Mateosian

2919 Forest Avenue

Berkeley, CA

94705-1310

(510) 540-7745

Best Science Writings—Readings and Insights, Robert Gannon, ed. (Oryx Press, Phoenix, 1991, 200 pp., $19.95)

Robert Gannon is a teacher of writing at Penn State. He spent 15 years as a free-lance science writer. He has assembled a small collection of award-winning examples of science writing. He says:

> This collection is designed mainly for people who want to write. It's for both beginning and established writers who would like to expand their area of interest, and it's for scientists who want to tell the world about new research.

Gannon spends a couple of pages introducing each of his 12 selections. He gives the backgrounds of the articles and the awards they won. These are quite interesting in their own right. For example, "Totality—a Report" by Michael Rogers appeared in *Rolling Stone*, not the usual source of award-winning science writing. Rogers wrote the article, about total solar eclipse addicts, for no money. He received only the expenses of a trip to Mauritania to watch an eclipse there.

Several of the articles have controversial aspects. For example, Carl Sagan's "Star Wars—the Leaky Shield" details his objections to President Reagan's Strategic Defense Initiative. Myron Levin's "The Greatest Health Protection in Cigarette History!" documents his investigative reporting of the effects on cigarette workers of the asbestos in Kent cigarette filters. Morton Hunt's "Research Through Deception" explores the ethical issues involved in psychological experimentation in which the subjects are deceived about the nature and purpose of the experiment.

All of the articles in this collection share one property. They are clear, fascinating presentations of difficult material. This is the essence of good science writing. If you're looking for models to help you improve your own writing, or if you just want to read a collection of interesting articles, get this book.

Database 101, Guy Kawasaki (Peachpit Press, Berkeley, 1992, 176 pp. plus disk, $18.95)

Where I went to school, admittedly a long time ago, the course designation "101" denoted an upper-division class. Kawasaki has written a book for fourth graders. Of course, the average fourth grader can easily understand everything that most people need to know about databases. I've found many expositions of database concepts to be unnecessarily abstruse and difficult, so Kawasaki has filled a real need.

Kawasaki is a former Macintosh evangelist. In 1986 he founded Acius, publisher of 4th Dimension, a large and extremely powerful relational database. In this book, however, he uses two simple nonrelational Macintosh databases for all of his examples. These are File Maker Pro from Claris Corp. and Touch Base from After Hours Software. The book comes with a disk containing "test drive" versions of these databases. As a reviewer, I can't use test drive versions, but you might find them interesting to play with if you're new to databases.

This very simple book covers all the main points. If you don't allow yourself to become offended at the baby-talk approach, you might find it a good introduction to databases. If you can't use it yourself, consider giving it to a fourth grader.

Borland C++ and Application Frameworks for IBM PC Compatibles, (Borland International, 12 manuals, 12 disks (5.25 in.), 10 disks

0272-1732/92/0800-0006$03.00 © 1992 IEEE

(3.5 in., high density), $749)

This is a truly monumental package. It took 45 Mbytes of hard-disk space to install it. As with Microsoft's similarly large C and Fortran packages, I couldn't get it running at first. Now, I've never been able to get the Microsoft packages working, but Borland's technical support got me started with one phone call. (The problem was a cache program in my autoexec.bat; it grabbed all of the available memory, preventing Borland's memory manager from initializing itself.)

Borland's package comes in a carrying case labeled with a warning about how heavy it is. The 12 manuals account for most of the weight. They describe the 45 Mbytes of software contained in compressed form on the disks. In addition to a C++ compiler, the package contains a C compiler, an assembler, a debugger, a profiler, libraries, complete support for Microsoft Windows, and an integrated programming environment, which includes an editor and a Make facility.

No doubt there will be feature wars and benchmark battles between Borland and Microsoft, but I don't intend to take sides. This brief review of the Borland offering is not a comparison with anyone else's package. If I ever get my Microsoft package running, I'll try to review it.

I have heard for many years about the integrated programming environments that were just around the corner. This one is definitely here, and it is head and shoulders above anything that I've ever used. Earlier programming environments were designed for assembly language programming; high-level languages were grafted on clumsily. Borland's package ties everything back to the original source code. It reports errors during compilation by highlighting the portion of the source statement giving rise to the error. During debugging, it allows stepping by source statements and displaying of expressions using source variable names. During execution, it allows the programmer to view simultaneously the source code, the output window, and dynamically updated variable traces.

A thorough review of this product would require far more time and space than I have available. I hope you can see from the little I've said how enthusiastic I am about it. If you are interested in experimenting with C++ to get a feel for what object-oriented programming is all about, this is an excellent way to do so. For not much money, you can obtain a professional programming tool that will let you put together simple programs or complex Microsoft Windows applications.

Borland C++ Handbook, 2nd ed., Chris H. Pappas and William H. Murray, III (Osborne/McGraw Hill, 1992, 966 pp., $29.95)

The folks at Osborne/McGraw Hill like to produce large books with large target audiences. In the preface to this one the authors promise:

> This book teaches you C, C++, and assembly language from the ground up and was designed for either the novice or the professional [and] those interested in learning more about procedure-oriented and object-oriented programming.

The saving grace of this tome is that it is small by comparison with the 9-inch stack of manuals that accompanies Borland's software. In fact, the book begins by orienting you to the manuals and other contents of the Borland package.

I used the chapter entitled "Getting Started with the Borland C++ Compiler" to tour the integrated programming environment, and it really helped, although I ran into some problems. The book asks you to enter manually a page-long C program, complete with errors that the authors have placed in it intentionally. Naturally, I introduced a few errors of my own unintentionally, so I quickly got off the path that the authors had planned. At other places, after I returned to the intended path, the Borland package functioned differently from the way the book seemed to say it was going to function. For example, I wasn't able to use watch windows by following the book's directions. By that time, however, I had become sufficiently oriented to the environment, so I stopped using the book's examples.

I think that I will continue to find this book useful as I work with the Borland package. However, I would not recommend it as a primer on C++ or C. Many other books designed specifically for that purpose are much shorter than this one.

Guest Editor's Introduction

Microelectronics in Europe

Egon Hoerbst

Siemens A.G.

For many years microelectronics has been one of the top discussion items of European managers, politicians, and scientists. In this issue we continue the dialogue with five articles from five different companies that describe the major European industrial activities in the various areas of microelectronics.

To describe the importance and the weakness of microelectronics in Europe, I will use the statements from JESSI's (the Joint European Submicron Silicon Initiative) 1990 Strategy Document.

- In the past Europe's microelectronics industry has succeeded in achieving a basic technological level matching that of the United States and Japan by means of extremely high efforts in research and development. However, Europe has hardly improved its market position. European-owned companies supply only about 40 percent of the European market for integrated circuits.
- Only in areas such as entertainment electronics or automotive electronics, in which close relations between IC manufacturers and users exist, could the potential of European-produced microelectronics in realizing competitive advantages be sufficiently exploited.
- The supply industry for equipment and materials needed for microelectronics production maintains only a very weak position in the market, despite the presence of many companies with considerable initiative and performance capabilities.

For Europe and the European industry at large, microelectronics is of vital importance. It is the most important base technology, and for a growing number of industrial branches its application determines the future innovation and competition potentials.

The electronics industry in Europe has reached the size of the automotive industry with a total value of about 70 billion ECUs (European counting units). The electronics industry is expected to grow at least twice as much as that of the average industry, making electronics in the year 2000 by far the biggest industry for Europe, with a size of about 150 billion ECUs.

Consequently, the availability and capability of a European microelectronics industry are essential for the technological infrastructure of the whole industry. The industry must perform two main functions

1) secure a reliable supply of electronic components, and
2) use to the full extent its potential for product and process innovation as well as for cost reduction.

In view of the fast technological progress of

0272-1732/92/0800-0008$03.00

microelectronics, the industry must participate in the most modern semiconductor technologies and rapidly convert them into advantages of product-performance and competitiveness in electronic equipment and systems.

The European microelectronics industry is subject to three basic problems:

- *Insufficient economy of scale.* The size of the individual microelectronics producers in Europe is insufficient to bear the expenditures for technological development and manufacturing investments, and the associated risks, which are required to keep up with the speed of technological innovation of their much larger competitors.
- *Insufficient vertical integration.* The layers of the electronics industry chain are heavily interdependent and long-term strategic advantages can only be achieved when all layers are available and provide state-of-the-art products. Increasing foreign penetration into the European industry chain, accomplished for instance by offering instant profits, decreases its strengths and creates undesirable dependencies.
- *Limited exploitation of available strengths.* Europe in general has strong capabilities in the different layers of the microelectronics industry but a weakness in exploitation. The traditional strength of the European industry in machine tools and precision mechanics should, for instance, be a healthy basis for semiconductor manufacturing equipment. Excellent capabilities exist in basic research at institutes and universities, but Europeans have achieved only limited impact on the industrial progress.

Because of the importance of the electronics industry, it is imperative that a strong electronics competence be retained in Europe. Therefore the European industry developed the JESSI strategy. Built around submicron silicon IC technology, in which many categories of the European electronics industry and academe play a part, JESSI aims to strengthen the total capability. The European industry is characterized by a structure in which medium-size electronics companies have a relatively high share in value adding and innovation. All initiatives will therefore take the specific requirements of this industrial structure into consideration.

JESSI's mission is to

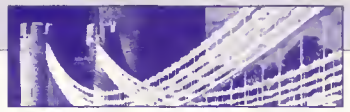
- strengthen the whole European electronics industry chain of the electronic systems, microelectronic components, and semiconductor production equipment and materials industries, and secure, also in the future, its worldwide competitiveness;
- secure the availability of European resources for the design, manufacturing, and application of microelectronic components, and stimulate the (vertical) relations between these resources; and
- stimulate the growth and market pull of the European markets for electronic systems, microelectronic components, and semiconductor production equipment and materials.

The five contributions in this issue describe a part of the work established under JESSI and ESPRIT (European Strategic Programme for Research in Information Technology) funding and give a good picture of today's work in Europe.

Egon Hoerbst is executive director at Siemens AG Corporate Research and Development and head of the System Technologies, Components and Devices Department. He is also a professor at the University of Munich and a guest professor at the University of Brunel, United Kingdom. His main activities comprise research in the fields of computer architecture, design automation, and ASIC cell libraries.

Hoerbst received his PhD in mathematics from the University of Innsbruck, Austria. He is a member of the International Federation of Information Processing, Gesellschaft fuer Informatik, Gesellschaft fur Mikroelectronik, and the Association of Computing Machinery.

Direct any questions regarding this issue to Egon Hoerbst, Siemens AG, ZFE ST KM, Otto-Hahn-Ring 6, D-8000 Munich 83, Germany.



CMOS Technology Trends and Economics

CMOS has become the mainstream IC technology. Extending well into the sub-0.1-µm regime, its potential provides enormous chip complexities for integration of complete systems on one chip. CMOS ICs by themselves will form a significant estimated market of US$150 billion in 2000. But CMOS is also an enabling technology for a large variety of high-technology industries with 1 or 2 orders of magnitude bigger markets. Therefore unrestricted availability of this technology is of strategic importance for the European high-technology industry. Exploding development costs and investments per technology generation require global cooperation, particularly for the relatively small European IC manufacturers to survive in this key technology.

Armin W. Wieder

Franz Neppl

Siemens AG

In the last decade CMOS evolved from a special technology for a few battery-driven devices and military applications into the mainstream technology for high-density VLSI circuits. By the year 2000 the current 55 percent of the total IC market for CMOS circuits will have increased to about 85 percent. CMOS then will clearly be the dominant technology for DRAMs (dynamic RAMs), SRAMs (static RAMs), EPROMs (erasable programmable ROMs), microprocessors, microcontrollers, complex gate arrays, and standard cell libraries for fast realization of customer-specific ICs, as well as for all types of application-specific ICs (ASICs).

The reason for this overwhelming success is that CMOS widely fulfills the needs of current and, even more important, future applications.

The powerful concept of scaling (the simultaneous reduction of all physical dimensions) provided a relatively straightforward way to significantly increase at the same time integration density and circuit speed from generation to generation. Figure 1 visualizes the progress of miniaturization made from the early 1970s to the 1980s. This trend is continuing at about the same rate. Several applications drive the practically unlimited need for increased integration density and speed. They are advanced RISC processors with huge on-chip cache-memories (Amdahl's law) and signal processors with extreme parallelization for high data rates, neural networks, robots, and speech recognition systems. (See Tables 1 and 2.)

The low power dissipation of CMOS circuits provides low power delay and allows very high integration densities without the penalty of excessive chip temperature or cooling cost. It also makes CMOS particularly suitable for ICs in the worldwide-proliferating pocket-size—and therefore battery-powered—systems.

The wide noise margin makes CMOS circuits relatively resistant to variations of supply voltage, temperature, and process. This robustness is also a good basis for the wide application of automated design techniques indispensable for handling the high complexity of future ICs and for fast realization of ASICs. Furthermore, it makes CMOS the technology of choice for the low-voltage, low-power operation of future ICs, particularly those used in mobile systems.

The price for these still quickly improving features had—at least in the past—been kept relatively moderate by the scaling inherent increase of the number of devices per area unit and per wafer. Another factor is the increasingly higher yield levels driven by high-volume commodity products such as DRAMs.

Finally, the successful demonstration of functional sub 0.1-µm CMOS devices[1,2] suggests the lack of any fundamental physical or technical limitations to CMOS technology, at least to about the 0.05-µm gate length. Such a CMOS technology would provide about 10^{10} devices per chip and

0272-1732/92/0800-0010$03.00 © 1992 IEEE

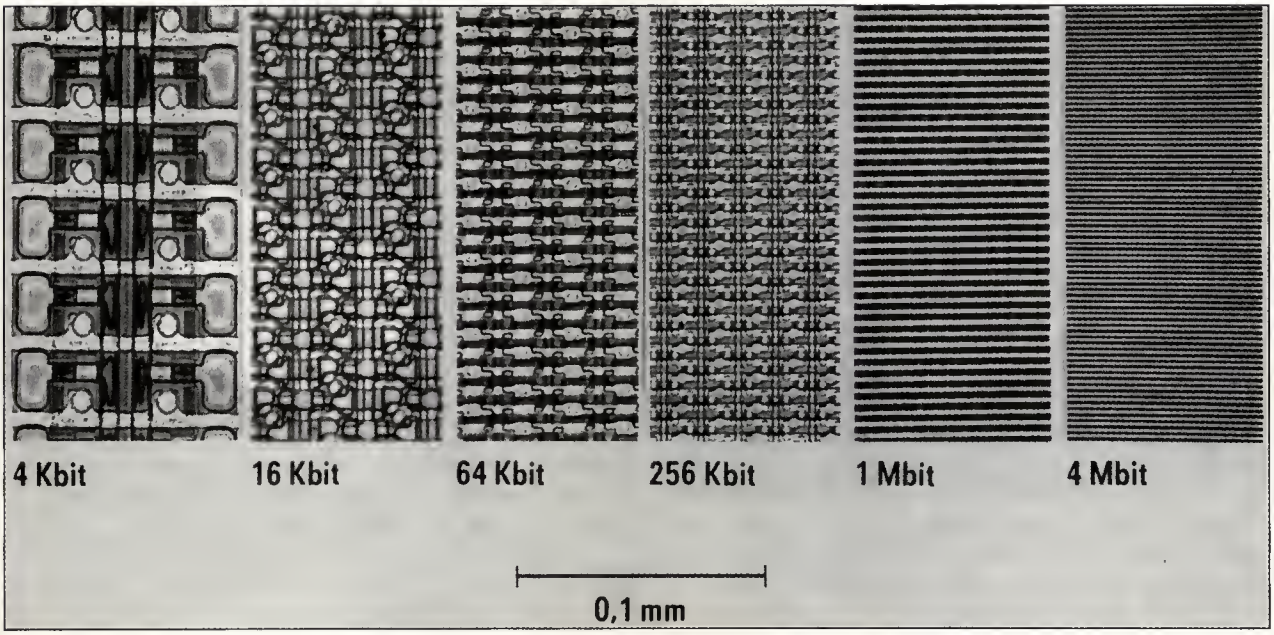


Figure 1. Reduction of feature size from a 4-Kbit to a 4-Mbit DRAM.

gate delays on the order of 10 picoseconds. Thus CMOS still has an enormous potential to be exploited in the future by IC and system designers, continuing to be the dominant IC technology for at least two more decades.

General trends

We can see a number of general trends in the scenario of future development and manufacturing of CMOS technologies and ICs.

The number of process steps to fabricate, for example, DRAMs, has increased from about 80 for the 16-Kbit version to about 450 for the 16-Mbit chip. The cost factor related to this steep increase of process complexity must be at least partly compensated by increased output. One trend based on this consideration is the move to larger silicon wafer diameters. Currently a change from a 6-inch to an 8-inch wafer diameter is taking place, and there are 12-inch concepts already. Larger wafer diameters, however, impose a significant challenge on wafer and equipment manufacturers with respect to material and process uniformity as well as to wafer handling. A tight interaction between equipment and material suppliers and IC manufacturers will be necessary to assure smooth transitions.

The high process complexity, the mutual interaction of process steps, the tight links between equipment and process, and the pressure for cost reduction make progresses achieved in research laboratory environments less and less significant. Only a concurrent development effort combining the skills of researchers and manufacturing specialists right from the beginning of development will ensure success in the future. Manufacturability

Table 1. Future requirements for speed in MIPS.

Application	Speed
Robotics	10^3
Expert systems	10^3-10^5
Speech recognition Vocabulary: 10^4 words, real time)	10^5
Image recognition (HDTV quality, real time)	10^6

Table 2. Future requirements for integration density.

Application	Requirements
Neural networks (parallel and associative information processing)	0.7-µm CMOS →10^3 neurons (mosquito: 10^3 neurons) 0.05-µm CMOS →10^5-10^6 neurons (human brain: 10^{12} neurons)

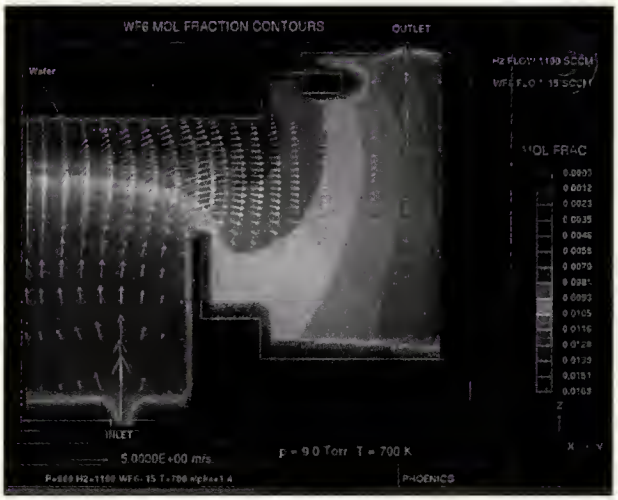


Figure 2. Optimization of tungsten-CVD deposition by simulation of gas concentration and gas flow.[3]

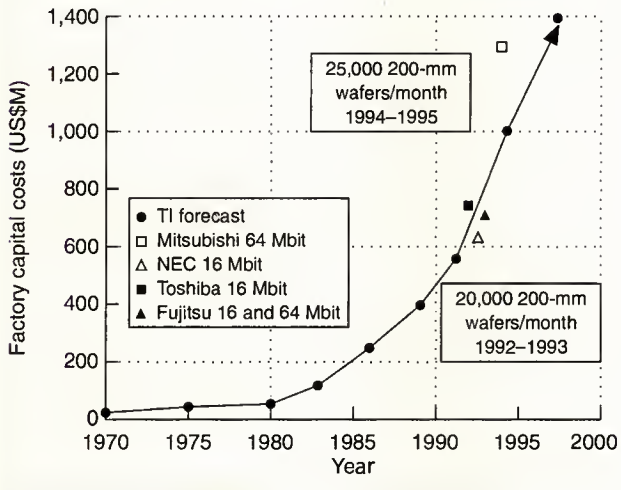


Figure 3. Factory capital cost trend.[4]

must be designed into the process from the beginning to support a smooth transfer into production and to provide the extremely high yield levels required. Again early involvement of equipment suppliers is indispensable, but additionally strong support is needed and expected from equipment and process simulation. Figure 2 shows an example of uniformity improvement of a chemical vapor deposition (CVD) system.

In view of the high process complexity, of the process interdependencies, and of the time and cost pressures, the variety of application-specific technology requirements can only be fulfilled in the future by a strictly modular technol-

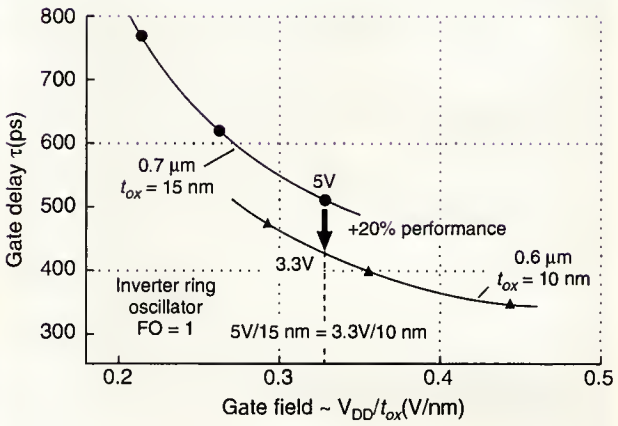


Figure 4. MOSFET optimization for optimum speed performance at reduced supply voltage.

ogy concept. This concept is based on a CMOS core process supplemented by various process modules for specific applications such as DRAM cells, load resistors or capacitors, nonvolatile memory arrays, or additional active devices like bipolar transistors. Preferably, a high-volume commodity product like the DRAM drives the quality of the core technology. Then, additional modules are subsequently introduced and evaluated in this mature core technology.

The predicted cost explosion of conventional IC fabrication (Figure 3) is caused by equipment cost, clean-room facilities, and a process complexity together with the need for tighter and preferably in-situ process control in the deep sub 0.5-µm regime. This will most likely enhance the search for revolutionary alternative fabrication concepts that are less costly and capital intensive. Cost-efficient manufacturing has to be supported by aggressively simplified process architectures with a significantly reduced number of process steps.

The long-preserved 5V power supply standard will finally be dropped due to device reliability and power consumption concerns. For the 0.5- to 0.3-µm CMOS generations, a new 3.3V standard was established, with further reductions of the supply voltage foreseen. Whereas this trend is beneficial for many applications—especially all battery-powered ICs for handheld, pocket-size systems—problems exist with analog circuits and with the interface to 5V circuits. The implementation of 5V circuitry and on-chip voltage reduction at least in a transition phase will mitigate these problems. To maintain a performance advantage over preceding CMOS generations despite the reduced supply voltage, device engineers will have to optimize the device structures regarding current drive and reliability. Figure 4 gives an example of such an optimization.

The enormous integration density of future CMOS generations of 10^9-10^{10} devices per chip in about 2010 will enhance

the trend toward increasing system integration. This will imply integration of complete circuit boards on one chip as well as, for example, completely new chip systems with extremely high data rate capability achieved by pronounced parallel processing. We see many chances for innovative system solutions. For example, the integration density of 10^{10} devices per chip would allow the implementation of a thousand 4-Mbit DRAMs, that is, 250,000 pages of information or 10^3-10^4 Intel 80486 microprocessors on one chip. Integration of logic and memory or even additional functions like optical components is very likely. Useful exploitation of this tremendous potential requires intensive interaction and cooperation between IC and system specialists and will favor system suppliers in developing such innovative "systems on chips." We expect to see extensive mutual stimulation between advanced CMOS IC technology and innovative system ideas.

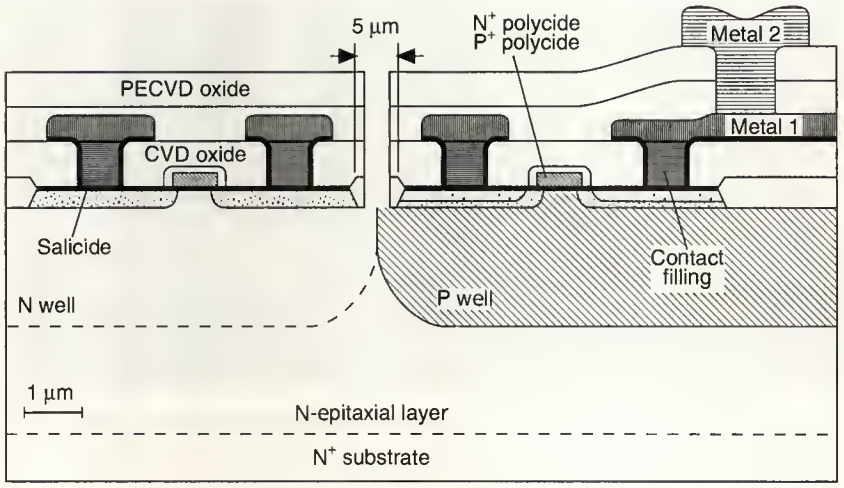


Figure 5. Schematic cross section of a 0.6-µm CMOS structure.

The high integration density requires a widely automated design technique based on gate array blocks, standard cells, and application-specific macros to guarantee sufficiently short development time for innovative ICs. This approach calls for a robust process and device architecture. Furthermore, however, the focus of technology development will have to be extended from MOSFET construction and optimization to high-performance interconnection systems with several levels of interconnections to handle the extreme wiring requirements of highly complex ICs.

CMOS technology today

Today's modern, state-of-the-art CMOS production processes feature 0.8- to 0.6-µm minimum design rules at least in some critical gate, isolation, contacts, or one-metal levels, depending on the product requirements. G-line or i-line, 5:1 optical stepper lithography and anisotropic dry etching are commonly used to produce chip patterns.

A typical process concept contains:

- P epitaxy on a highly doped P+ substrate;
- a double-well structure for sufficient latch-up resistance;
- advanced Locos (local oxidation of silicon isolation) with bird's beak (tapered oxide edge) lengths reduced to 0.1 µm;
- 10- to 15-nm silicon dioxide as a gate dielectric;
- N+-poly- or polycide gate;
- a "lightly doped drain" structure optimized for sufficient reduction of hot carrier effects and minimum series resistances, that is, the highest possible current drive of the NMOSFETs; and
- two to three levels of metal interconnections with tungsten-filled contacts for high-packing density and current capability.

Figure 5 displays a typical cross section.

The internal supply voltage is commonly 3.3V. Since this CMOS generation is characterized by a transition of supply voltage from 5V to 3.3V however, 5V-compatible input/output circuitry with on-chip voltage reduction is also provided to mitigate interface problems on system boards.

The leading product of this CMOS generation is the 16-Mbit DRAM. Figure 6 (on the next page) shows a microphotograph of the chip. Siemens, the only European supplier with competitive engineering samples, based its timely and proprietary development on the solid development and manufacturing know-how acquired with its 1-Mbit and 4-Mbit products. For economic reasons Siemens cooperated with IBM in France on the 16-Mbit fabrication. The arrangement will provide a worldwide competitive, state-of-the-art 0.6-µm CMOS fabrication capability in Europe by 1992. Two other major European IC suppliers are developing 0.6-µm CMOS technology (SGS-Thomson mainly for EPROMs and Philips until recently for fast 4-Mbit SRAMs). Even so, the Japanese, US, and Korean manufacturers largely dominate the IC market in submicron CMOS technology.

Future trends

Predominantly driven by the DRAM market, a large amount of research and development effort is dedicated to the investigation and evaluation of all aspects of future CMOS tech-

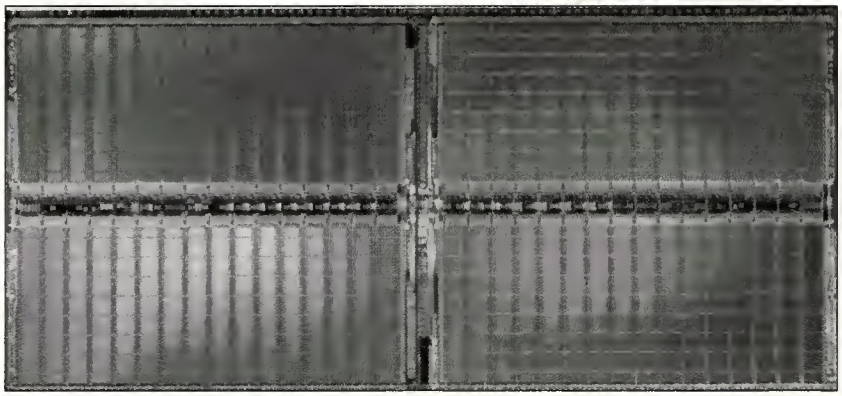
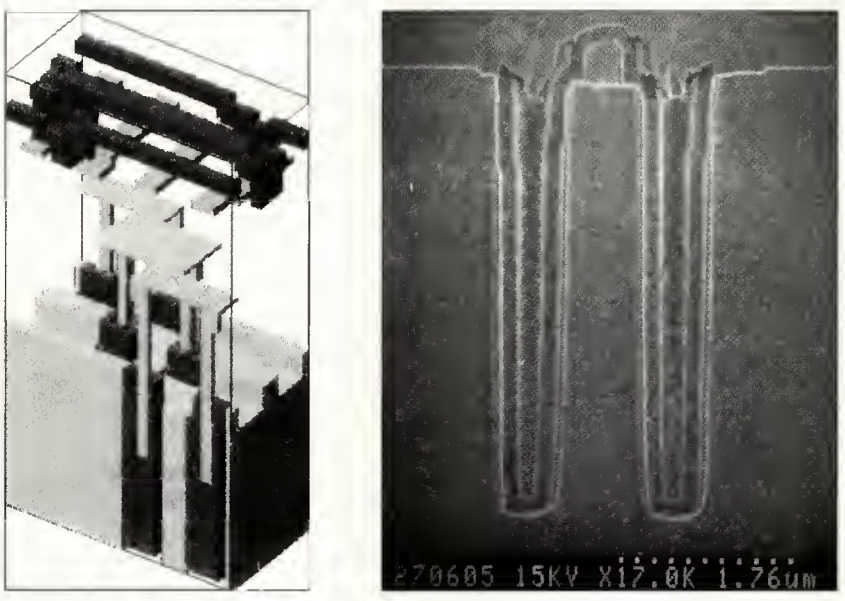


Figure 6. A 16-Mbit DRAM processed with 0.6-µm CMOS; the cell size is 1.6×3.0 µm^2, and the chip size is 144.7 mm^2 for 33.5 million devices.



(a) (b)

Figure 7. SEM cross section (a) and schematic 3D view (b) of a 64-Mbit stacked-trench-capacitor cell.[5]

ning electron microscope view of a 64-Mbit-capable trench capacitor cell. The technology straightforwardly extends the 16-Mbit concept. Product development continues in all major semiconductor companies, with market introduction scheduled for 1994.

Concepts for the subsequent 256-Mbit DRAM generation are widely published. Its feasibility is unquestioned today, with market introduction planned before the year 2000.

From the results of exploratory work we can conclude that the relatively conventional CMOS device concepts are working down to at least 0.05-µm gate lengths. This leaves room for a number of additional CMOS technology generations well into the gigabit regime.

Some of the essential innovations in the different specific areas of CMOS are highlighted here.

Based on the scaling principle for improved integration density and circuit speed, progress in CMOS technology is tightly linked to the progress in lithography necessary to print the further and further reduced lateral dimensions.

Recent results obtained with optical i-line lithography in combination with the advanced "phase shift" mask technique[6] demonstrated the possibility of optically printed structures below 0.2 µm. Here, creating and superimposing light patterns with a 180-degree phase difference achieves enhanced edge contrast. This is far below the originally assumed limit of optical lithography and suggests the applicability of i-line lithography for the 64-Mbit and maybe even the 256-Mbit generation. Extension of this technique to still shorter wavelengths of the light source (deep ultraviolet) and the use of holographic principles will take optical lithography into the gigabit regime. Several laboratories also in Europe currently develop deep UV lithography.

nologies. As in the past, Europeans will make some significant contributions to this field. Whether these will be enough to survive in such a highly competitive market is doubtful, however.

A clear technology road map to the next CMOS generation is visible today. The leading-edge product, the 64-Mbit DRAM, will offer about a 35-ns access time, a 0.35-µm minimum feature size, and a 180-cm^2 chip size. Figure 7 shows a scan-

Advanced multilayer resist technologies such as Carl[7] and Desire[8] with top-surface imaging support the progress in optical lithography. The problems of resist thickness variations and depth of focus limitations are eliminated and contrast is enhanced. Figure 8 demonstrates 0.25-µm lines using the Carl technique.

To make these advances in lithography usable in produc-

tion environments, however, simultaneous improvements of level-to-level alignment accuracy must be achieved.

For a while X-ray lithography was widely considered to replace optical lithography in the deep submicron regime and was even a candidate for 64-Mbit DRAM production. As a consequence of the improvements in optical lithography, however, introduction of X-ray lithography is likely only below a 0.15-μm minimum feature size. There, X-ray lithography, which is still plagued by severe mask problems and a high investment barrier (US$40 million per lithography unit), most likely will encounter strong competition by advanced electronic-beam and ion-beam systems.

In CMOS device isolation there will eventually be a change from the conventional implanted-well approach with long high-temperature drive-in cycles toward retrograde-well structures. These latter structures provide lower well resistance and therefore better latch-up hardness even at very small spacings between N- and P-channel MOSFETs and without the penalty of deteriorated device characteristics and junction capacitances. Elimination of the long, high-temperature drive-in annealing makes the process more compatible with the trend to single-wafer processing. The retrograde-well technique with relatively high implantation energies also permits the implantation of the wells after the isolation technique and thus automatically establishes a high field doping. Both aspects support future manufacturing concepts and the trend to simplified process architectures.

Even in the 64-Mbit generation, isolation of MOSFETs inside one well is mostly ensured by advanced modified Locos techniques, most likely because of the low defect levels achieved in comparison with shallow-trench isolation concepts. The tapered oxide areas (bird's beaks) with reduced isolation oxide thickness, which are inefficient for device isolation, could be reduced to below 0.1 μm. Only when isolation spacings drop below 0.4 to 0.3 μm will shallow-trench isolation be widely adopted. The ultimate solution for both types of device isolation may be the silicon-on-insulator (SOI) approach.

Successful scaling of the N- and PMOSFETs into the deep sub 0.5-μm regime has several aspects. Extremely thin gate dielectrics are required. Reliable, defect-free silicon dioxide films as thin as 4 nm have been demonstrated. This is close to the direct tunneling limit of about 3 nm. Enhanced conformity, relative insusceptibility against surface contamination (reduced pinhole density), and reduced dopant segregation favor rapid thermal oxidation at relatively high temperature. Nitridation of the silicon dioxide gate dielectric could further enhance reliability and yield.

Whereas the trend of channel doping for deep sub 0.5-μm devices is straightforward for the surface-channel NMOSFET, the situation of the buried-channel PMOSFET with a compensated surface for sufficiently low threshold voltage is more complicated. One report provides a thorough investigation of the buried MOSFET.[10] Even though buried-channel

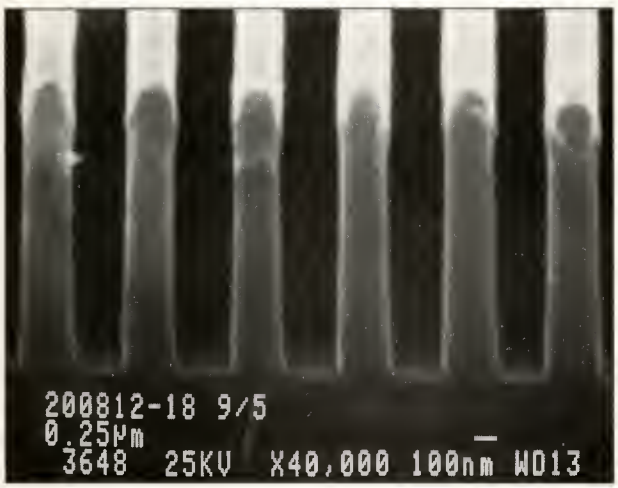


Figure 8. 0.25-μm resist structures using the Carl technique.[9]

PMOSFETs with a gate length of 0.1 μm have been demonstrated, a transition to the dual work function technique with P+-polycide gate PMOSFETs appears to be more likely. The work function of P+-polycide gates will allow a surface-channel construction for the PMOSFET and eliminate the need for dopant compensation at the surface for low threshold voltages. Figure 9 on the next page gives an example. Device construction and scalability would become widely comparable to the less critical NMOSFET. Process architecture could become simpler, for example, by eliminating all channel implants, corresponding annealing, oxidation and etching steps, and photomasks. The lower threshold voltage would overcompensate for the reduced hole mobility of the surface channel, particularly at further reduced supply voltages. Appropriate processing and/or gate oxide nitridation will suppress the deteriorating boron penetration through the gate dielectric.

Simple scaling theory asks for extremely shallow source and drain-junctions with $xj < 0.1$ μm for sufficiently good short-channel behavior in the deep sub 0.5-μm regime. Care must be taken, however, to find an acceptable compromise of the profile design for hot carrier injection and parasitic series resistances. Preamorphization by high-dose implantation of silicon or germanium to suppress channeling and reduce diffusion has been widely studied. Since silicidation of the relatively high-resistive source/drain areas will become standard for many applications, the technique of silicide formation, source and drain-implantation into the silicide, and subsequent annealing for dopant out diffusion seems to be more promising.

Figure 10 shows a result of an IMEC/Siemens cooperation demonstrating P+ junctions shallower than 0.1 μm. This junction formation is self-aligned to the metal/silicon interface and thus guarantees good diode quality even at extremely low

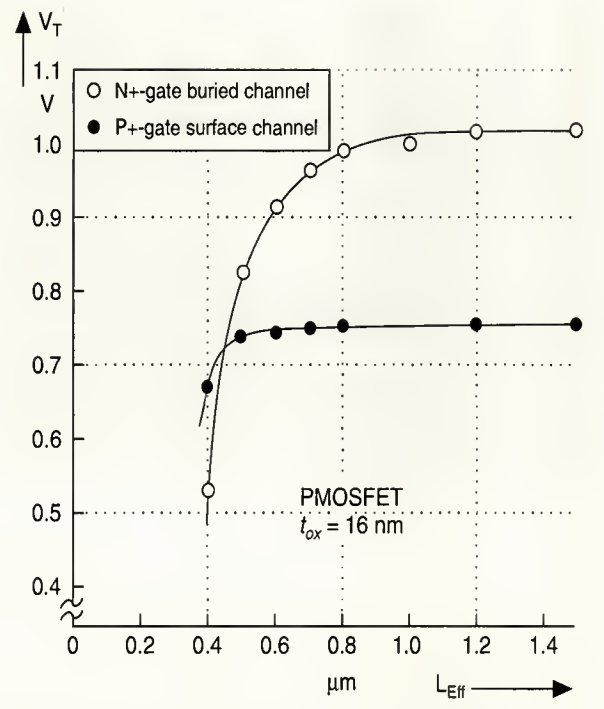


Figure 9. Improvement of PMOS short-channel characteristics by work function engineering.

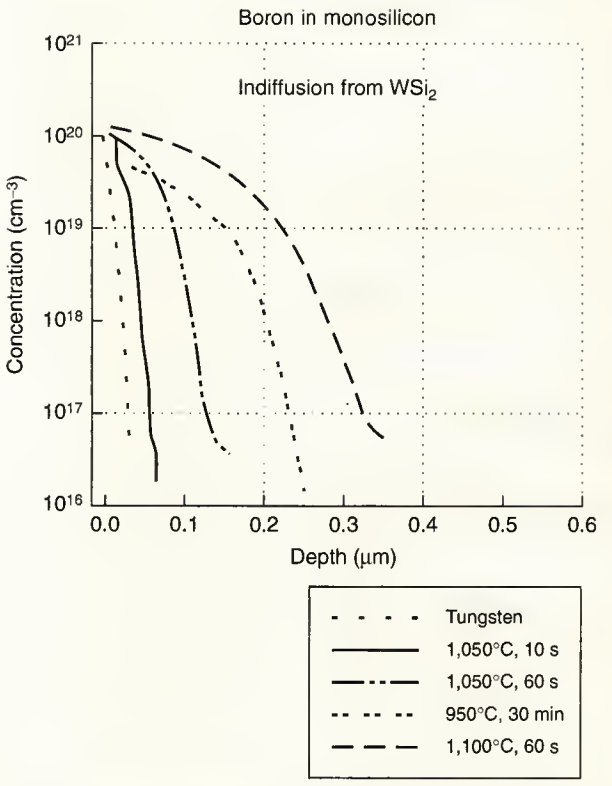


Figure 10. SIMS (secondary ion mass spectroscopy) depth profiles of boron indiffusion from WSi_2 into monosilicon after different rapid thermal annealing (RTA) and furnace annealing (FA) cycles.[11]

junction depths and relatively rough silicider/silicon-interfaces.

Finally, the voltage reduction, more realistic life time criteria with respect to hot carrier degradation,[12] and the results of a thorough investigation of the scaling of the hot carrier-induced device degradation indicate enough reliability margin for designing high-speed, deep sub 0.5-µm CMOS FETs. Fundamental limits for the conventional MOSFET devices are not foreseen down to below 0.1-µm gate length.

Ultimately MOSFETs might rely on fully depleted structures or might contain δ-doping layers for improved short-channel characteristics.

As mentioned, availability of highly reliable low-resistive and dense interconnection systems with low capacitances is crucial for future highly complex ICs. A global planarization approach has to be implemented to eliminate topography-related problems (step coverage, depth of focus limitations) and to allow for a modular, multilevel metallization concept. Process complexity and difficulties per interconnection level would then no longer increase and change with the number of interconnection levels. This is a precondition for economical implementation of an increasing number of metal levels.

Enhanced requirements for reliability, defect density, and resistivity will change the current standard aluminum-silicon-copper alloy to new alloys, to multisandwich systems like aluminum/titanium nitride, or even to copper or gold. The commonly used tungsten contact plugs should be replaced by lower resistive material, preferably the interconnection material itself, to eliminate material discontinuities and temperature inhomogenities that usually enhance electromigration-induced failures. CVD metallization or the recently developed ultra high vacuum-sputtering technique (Figure 11) might be solutions.

Development of stable intermetal dielectrics with low dielectric constants would be highly appreciated for reduced interconnection capacitances. An increasing number of interconnection levels will increase the integration density of usable devices and additionally will reduce average capacitive load.

The uppermost metal level and the passivation must accommodate the special mounting and packaging techniques of future systems like advanced multichip modules.

BiCMOS (bipolar CMOS) technology combines all the CMOS features and advantages with the superior driving capability and the inherently better analog characteristics of bipolar transistors. BiCMOS might never really replace CMOS as the main-

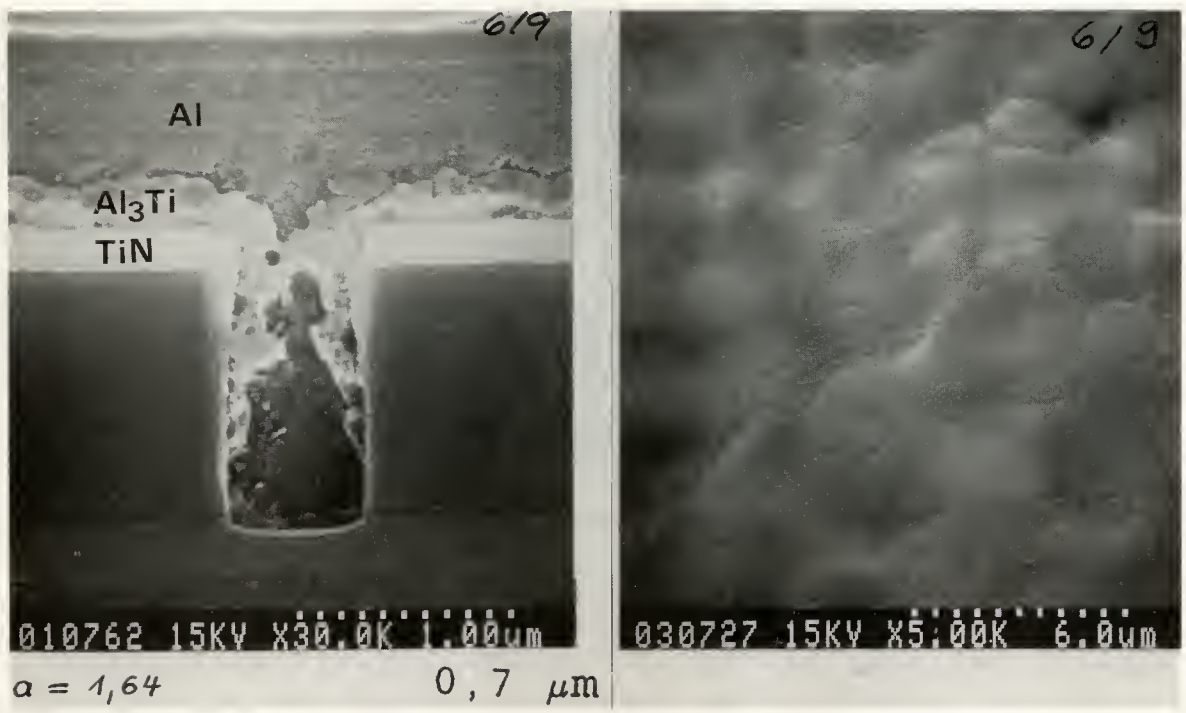


Figure 11. SEM cross section of a 0.7-µm contact hole completely filled and planarized by ultra high vacuum-sputtered aluminum (a) and a corresponding SEM surface image (b) of an array of 20 contacts.[13]

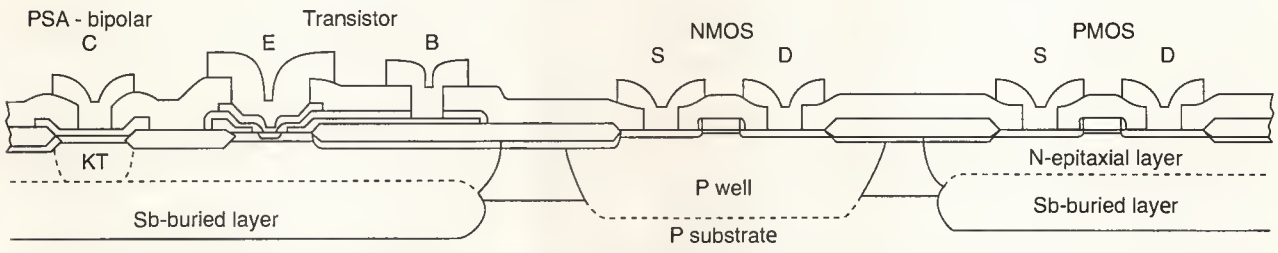


Figure 12. Schematic cross section of a 0.8-µm BiCMOS structure.

stream IC technology because of a 30 to 50 percent increased process complexity. Yet, it certainly will become an indispensable companion of CMOS, providing excellent on-chip analog functions, boosting performance for special applications such as fast cache memories, and finally providing better circuit performance at a given minimum feature size. The latter might extend the use of a given fabrication environment for still-competitive products and thus help to increase profitability. Figure 12 shows a cross section of a modern BiCMOS structure. Modular process architecture with CMOS is mandatory for cost efficiency.

As mentioned earlier, the ultimate CMOS solution might be a silicon-on-insulator approach. Advantages are the simplicity of the process architecture, improved short-channel characteristics of the fully depleted MOSFET structure, reduced power consumption, and reduced gate delay. A precondition for such a scenario is a successful technique to economically provide a high volume of silicon-on-insulator substrates. The wafer-bonding technique in combination with tightly controlled polishing of the top wafer to a thickness of the order of 0.1 µm might be a promising candidate.

Because of the wide variety of CMOS applications in pocket-size systems, low-temperature operation of CMOS (77 degrees Kelvin) with its speed advantage of a factor of about 2 will not become general practice for CMOS ICs. The size and cost of respective cooling systems will restrict this to medium and

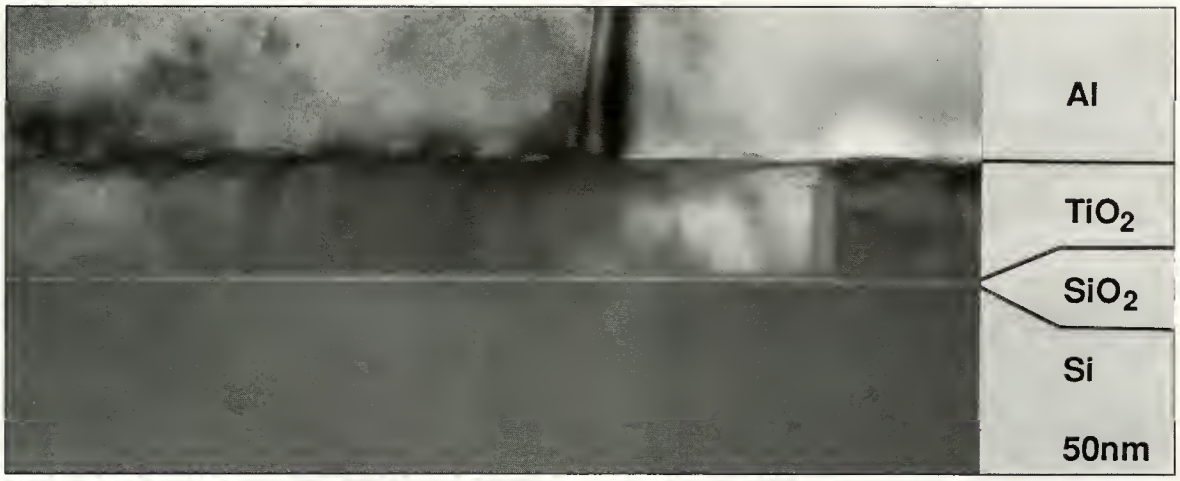


Figure 13. Cross-sectional TEM (transmission electron microscope) microphotograph of a 50-nm TiO_2 layer (effective thickness 2.5 nm) annealed at 800°C in Ar/O_2. Evaporated aluminum used as the gate electrode of the capacitor covers the dielectric.[14]

large-size systems.

The technology trends discussed so far refer to the CMOS core process. Of all the additional modules the storage cell for dynamic memories is the most prominent because of the leading role of DRAMs as the technology driver in each CMOS generation. One trend in DRAM cells is the introduction of innovative high ε dielectrics in the storage capacitor, such as Ta_2O_5 ($\varepsilon \approx 20$), TiO_2 ($\varepsilon \approx 100$—see Figure 13), or PZT ($\varepsilon > 1{,}000$). Such a material innovation would allow the extension of relatively simple and well-known cell structures such as the stacked-cell concept far into the gigabit regime.

Also, recent publications demonstrate that an additional potential still exists for more clever memory cell design to significantly reduce the cell area needed for a given capacitance. Finally, an intracell read current amplification by integrating a bipolar transistor into the bit-line contact hole showed the possibility of minimizing the cell capacitance to the soft error rate limited value.[15] All these results indicate that no fundamental technical barriers exist well into the gigabit DRAM regime.

CMOS: The economical scenario

While the results of a great deal of exploratory work suggest a technically unlimited potential for CMOS technology at least until the second decade of the next century, the possibility of economical industrial exploitation is not yet as clear. The usefulness of many of the concepts under high-volume production conditions must still be assessed. Near-exponential cost explosions in the corresponding developments and fabrication plants (see Figure 3 again) in combination with high entrepreneurial risk could soon surmount the financial power of even the biggest companies. Unrestricted application of all possibilities of CMOS technologies in the future will strongly depend on how fast and successful industry is in inventing and implementing more cost-efficient and less capital-intensive fabrication and development scenarios that will allow IC production in affordable financial dimensions.

CMOS technology and the corresponding ICs are a significant economic factor with many business opportunities. By the year 2000 the market volume will have increased to about US$150 billion. While Europe represents roughly 20 percent of the world market for ICs, European suppliers only account for a combined 10 percent market share. Japanese and American IC manufacturers, with shares of 50 and 35 percent, dominate the market. Europe imports more than 50 percent of its ICs—equivalent to US$18 billion in 2000. This fact considerably contributes to trade deficits, particularly with Japan. European suppliers must achieve a stronger position—at least in their home market—for more balanced trade.

Since ICs are key components for all modern telecommunication, automation, medical instruments, data, information, and traffic systems, CMOS is the most prominent enabling technology for industries with one to two orders of magnitude larger market volume than the direct IC market itself. Microelectronics will be an increasing part of all kinds of applications. Thus unrestricted access to leading-edge CMOS technology is not only essential for the survival of the European IC manufacturers but is of crucial strategic importance in maintaining the international competitiveness of a wide range of European high-technology industries in the future.

The dominance of Japanese IC suppliers is particularly critical, since they supply systems, too. They are in a position to use the dominance in a key technology like CMOS as a strategic weapon to gain dominance in all electronics-related systems markets.

The unsatisfactory position of the European IC industries for the coming challenges becomes particularly clear if one takes into account one fact. Development costs for the future CMOS technology generation and corresponding economical fabrication plants each generate about $US1 billion a year. For the leading Japanese companies with current total revenues from ICs of about $US6 billion each, such a sum amounts to about 15 percent of the revenue and therefore might still be affordable. For the European suppliers with total IC sales around $US1 billion each, however, this is 100 percent of the total revenue and clearly surmounts the economic power of each company.

IN VIEW OF THESE ECONOMIC FACTORS, cooperation strategies become vital for the survival of European IC manufacturers, which could in turn guarantee the unrestricted European access to the key technology of CMOS. Corresponding initiatives go back to the beginning of the 1980s, starting with the inter-European Mega project between Philips and Siemens that aimed at 1Mbit- and 4-Mbit DRAMs and 1-Mbit SRAMs. With the volume production of 1-Mbit and 4-Mbit DRAMs in Germany, Siemens established a worldwide competitive submicron CMOS production facility in Europe.

The JESSI project, including also SGS-Thomson and many renowned European research institutes, covers 0.7- to 0.3-µm CMOS technology for logic applications and exploratory work on 0.25-µm CMOS technology. JESSI also tries to promote CMOS applications and develop a stronger European equipment industry for IC manufacturing. This European cooperation scenario has to be supplemented with strong intercontinental cooperations to implement the world's best manufacturing know-how, to gain time-to-market advantages by cooperation with worldwide technology leaders, and to assure access to the two biggest IC markets (the US and Japan). Last but not least, such cooperation will reduce the financial risk to an acceptable level. Examples are the Siemens and Toshiba cooperation for 1-Mbit DRAM manufacturing and the joint development of the 64-Mbit DRAM by Siemens and IBM.

Both cooperation strategies have to be followed and coordinated to successfully establish a worldwide competitive IC industry in Europe.

References

1. G.A. Sai-Halasz et al., *IEDM Tech. Digest*, 1987, p. 397.
2. M. Aoki et al., *IEDM Tech. Digest*, 1990, p. 939.
3. C. Werner, *J. Physique IV*, Vol. C, No. 2 and 3, 1991.
4. G.B. Larrabes, *Proc. IEEE Int'l Semiconductor Manufacturing Science Symp.*, 1991, p. 30.
5. L. Risch et al., *Proc. SSDM 90*, Vol. 837, 1990.
6. C. Nolscher, *VDI/CDE Fachtagung, Maskentechnik fur Mikroelektronikbausteine, (Mask Technology for ICs)*, Nov. 14, 1991, Munich.
7. M. Sebald et al., *Microelectr. Eng.*, Vol. 11, 1990, pp. 531-534.
8. F. Coopmans and B. Roland, *Proc. SPIE*, Vol. 631, 1986, p. 34.
9. R. Sezi et al., *Jap. J. Appl. Physics*, Vol. 30, No. 11B, 1991, p. 3108.
10. F.M. Klaassen et al., *IEDM Tech. Digest*, 1984, p. 613.
11. V. Probst et al., *J. Appl. Phys.*, Vol. 70, No. 2, 1991, p. 708.
12. J. Winnerl et al., *IEDM Tech. Digest*, 1988, p. 204.
13. To be published in *Semiconductor Int'l.*
14. A.Spitzer et al., Proc. INFOS, 1991, p. 18.
15. S. Sunouchi et al., *IEDM Tech. Digest*, 1991, p. 465.

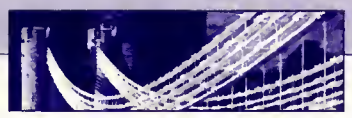
Armin W. Weider, as senior director, heads the silicon processing activities at Siemens Corporate Research and Development in Munich. He works on device, process, and circuit simulation; bipolar, CMOS, and BiCMOS circuit design; and process architecture. He has made contributions in the areas of device physics, exploratory bipolar technology, submicron CMOS and BiCMOS technology and circuitry, and the general area of the evolution of microelectronics to ever-smaller dimensions.

Wieder received the BS, MS, and PhD degrees in electronic engineering from the University of Aachen. He is a senior member of the Institute of Electrical and Electronics Engineers, has published more than 40 papers, and holds more than 20 patents in microelectronics.

Franz Neppl is deputy director of the Corporate Research and Development Department for silicon process integration including CMOS, BiCMOS, bipolar, and smart power technology. Since 1985 he has held responsibility for the development of submicron CMOS logic technology and exploratory work on 64-Mbit DRAMs. He has also worked on the integration of refractory metal silicides in NMOS and CMOS technologies, reliability of interconnections, and diffusion barriers.

Neppl received the diploma degree in physics and PhD rer. nat. from Technical University, Munich. He has published about 40 papers and contributions to international conferences and holds 35 patents in microelectronics.

Direct questions concerning this article to F. Neppl, Siemens AG ZFT BT ACM 1, Postfach 83 09 53, D-8000 Munich 83, Germany.



Consumer Applications:

A Driving Force for High-Level Synthesis of Signal-Processing Architectures

A growing need exists for digital signal processing in consumer applications. This overview of design alternatives for the translation of signal processing systems into silicon addresses that need. Since the most appropriate approach varies with the life-cycle phase of the application, the alternatives range from general-purpose to application-specific approaches. In particular, we review our achievements over the past five years in the field of high-level architecture synthesis.

Rob Woudsma

Jef L. van Meerbergen

Philips Research Laboratories

The continuous trend toward further inclusion of digital signal processing systems in consumer applications is exemplified by the compact disk—the most successful pioneering example[1]—and the digital mobile radio.[2] A variety of other examples yet to enter the consumer market will extend this trend: high-definition television,[3] digital compact cassettes,[4] and digital audio broadcasting.[5] The opportunities offered by state-of-the-art CMOS fabrication technology and the synergy of digital signal processing techniques and VLSI design triggered the development of these and similar products.

Such development presents system and semiconductor industries with a challenging opportunity to boost the application of VLSI systems and circuits. In fact, European companies with an established competence in signal processing systems may be particularly adept at meeting this challenge. Their competitive edge also gives them the potential to redress the imbalance caused by the leading position of Japanese VLSI technology. Pan-European research programs such as ESPRIT and JESSI strongly emphasize this consideration.

The advent of digital signal processing in consumer applications requires VLSI circuits that are commercially feasible in high volumes. Hence the primary design objective is to minimize hardware cost. Designers therefore will likely adopt an application-specific or full-custom design approach to minimize chip area. However, the associated design time could conflict with the short time-to-market requirement for consumer ICs. In recent years product lifetimes have been drastically reduced and in some cases are little more than a few months. Existing programmable digital signal processors (DSPs) are an important class of circuits in this context. They combine flexibility in functions and features with a short design time and a moderate development effort. Moreover, designers can easily realize derived products by modifying microcode.

A DSP must meet a well-defined speed-performance constraint that directly relates to the sampling rate associated with the application. This constraint contrasts with general data processing in microprocessor applications in which speed is an objective rather than a constraint. DSPs and other microprogrammed processors correspond to the extent that they permit sharing of hardware resources, merging of processing functions, and executing of control functions.

In the past decade numerous general-purpose DSPs became commercially available for a wide spectrum of applications. Lee gives an overview.[6,7]

0272-1732/92/0800-0020$03.00 © 1992 IEEE

These processors include fixed-point DSPs such as the Texas Instruments TMS320 series, NEC μPD77 series, Philips SP50 family,[8] Analog Devices ADSP-2100 series, SGS/Thompson ST18930, AT&T DSP series, and Motorola DSP56000. Floating-point DSPs such as the TMS320C30 and Motorola DSP96000 also achieved wide use.

All these processors combine the advantages of high application flexibility, immediate availability, and strong software and hardware support. Existing powerful DSPs offer extensive instruction sets, high levels of concurrency, large amounts of on-chip memory, and big off-chip memory address spaces. A true general-purpose signal processor, however, does not exist. Although video signal processors such as the VSP[9] have been presented, most general-purpose DSPs suit applications in areas with relatively low sampling rates such as digital audio, telecommunications, or speech processing.

Due to their general-purpose character these processors often display overhead in accuracy, memory size, or instruction set. Furthermore, application-specific operations sometimes execute very inefficiently. Often they also exhibit I/O limitations when communicating with other parts of a signal processing system. Consequently, general-purpose processors are usually either too costly or functionally inadequate for high-volume consumer applications. These processors are therefore often reserved for prototype or pilot versions in the first phase of the product life cycle.

To reduce these limitations, designers have introduced the concept of a DSP core, an architecture tuned to a certain application area. Typically integrated in combination with application-specific blocks on a single IC, the mask-programmable core offers flexibility at a lower price than a standard DSP. We can distinguish two forms of DSP cores. One is a fixed core derived from an existing general-purpose DSP and integrated with specific blocks to solve the I/O limitation problem. The second form is reconfigurable and optimizes a DSP by tuning a set of architecture parameters.[10]

The motivation behind the use of a DSP core is to trade hardware efficiency against development cost. For high-volume applications with a fixed functionality, however, even a DSP core may yield insufficient hardware efficiency. Consequently, designers need application-specific processor architectures that are fine-tuned to particular applications in terms of data path configuration, word lengths, and memory dimensions. Matterne et al. give an example.[11] Like DSP cores, these processors are usually embedded with application-specific modules like data or control interfaces, and produce an optimal performance/area ratio. Flexibility with respect to functional modifications is generally limited to pin-programmability, development times are long, and associated costs are high. Application-specific DSPs are apparently only acceptable in high-volume DSP products with more or less frozen functionality.

High-level synthesis

In an attempt to reduce development times and costs of application-specific hardware, researchers have introduced high-level architecture synthesis techniques. These techniques let designers evaluate various architecture alternatives using accurate feedback from the underlying libraries and layout environment. Even system designers who are not IC experts can perform this iteration process. Furthermore, automatic synthesis techniques contribute to proven correctness and testability of the resulting IC.

High-level synthesis boosts the number of application-specific solutions for DSP applications as development costs and turnaround times are drastically reduced. Numerous DSP compilers implementing these high-level synthesis techniques have been presented, mostly from universities and research institutes. While various high-level synthesis techniques and strategies were applied in these compilers, we can divide the approaches found in literature into two categories. The first includes general approaches such as Mimola,[12] Hal,[13] Maha,[14] Chippe,[15] Flamel,[16] Yorktown Silicon Compiler,[17] Easy,[18] System Architect's Workbench,[19] Caddy,[20] and Callas.[21] The second category includes domain-specific approaches such as First,[22] Lager,[23] Spaid,[24] Cathedral,[25] Piramid,[26,27] Hyper,[28] and Phideo.[29]

Following the concept of application-driven synthesis, the design of a DSP compiler starts from an analysis of the application field. A defined target architecture reflects the typical properties of the application domain that are exploited by the designers. This architecture model or architectural style defines the boundaries of the design space. We are convinced that a DSP compiler can only yield efficient results if it is based on an underlying target architecture model. The synthesis tools constituting the compiler can then optimize this architecture model to a given application. Using an architecture-driven approach means that the application area is restricted to those applications that fit within the architecture model. Hence we have defined a number of architecture models for complementary application domains and describe two representative models and associated compilers here: Piramid and Phideo.

The Piramid compiler

Piramid[27] supports digital audio, low-end digital video, telecommunications, speech processing, and control applications. For these DSP applications sampling rates are two or three orders of magnitude smaller than the clock frequencies that are feasible in today's CMOS fabrication processes. Consequently, sampling rates are typically limited to about 100 kHz at clock frequencies in the range of 10 to 25 MHz. Under this condition, designers can apply a highly multiplexed data path controlled by a central controller. Such a microcode-based approach can also be applied for complex decision-making applications.

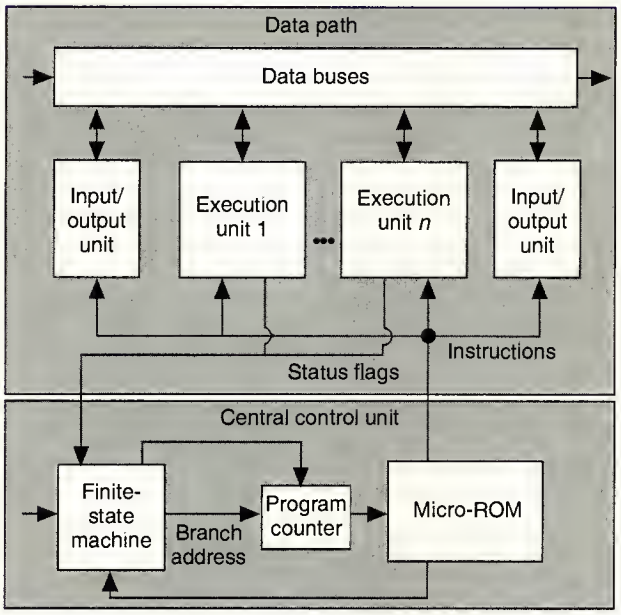


Figure 1. Piramid architecture model.

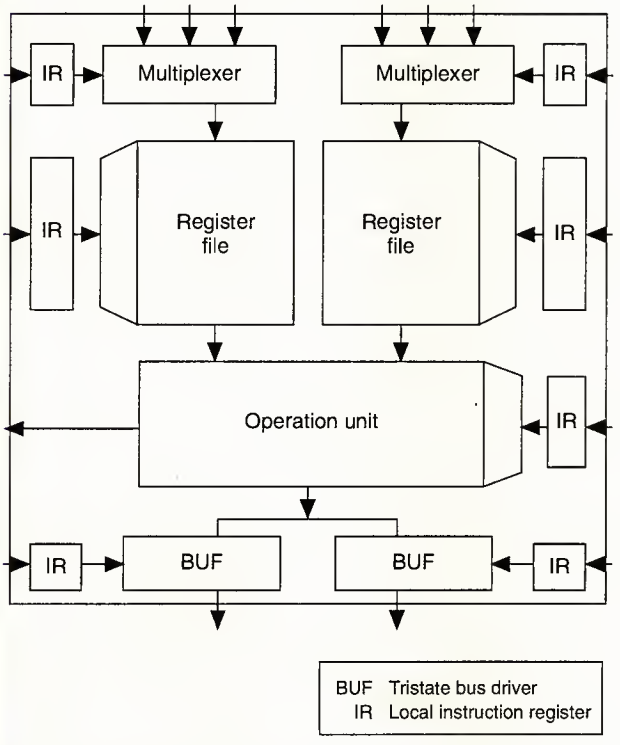


Figure 2. Execution unit template and types.

These architectures contain a limited number of arithmetic units and a small number of on-chip memories.

Architecture model. Figure 1 depicts the architecture model used by the Piramid compiler. It consists of a reconfigurable data path and a central control unit. The data path is an assembly of general-purpose processor blocks called execution units, which are connected by a number of data buses. For communication with other on-chip processors or with the external world, Piramid includes a number of input/output units in the data path. The Piramid control unit[30] consists of a microinstruction memory (micro-ROM), program counter, and multibranch finite-state machine. In every machine cycle, a microinstruction transfers from the micro-ROM to the execution units and the input/output units. The execution units generate status flags, which are stored in the branch finite-state machine to permit condition evaluation. All execution units can be simultaneously active, and the controller can evaluate multiple branch conditions within the same cycle. Hence the potential parallel processing power of this architecture can be fully exploited.

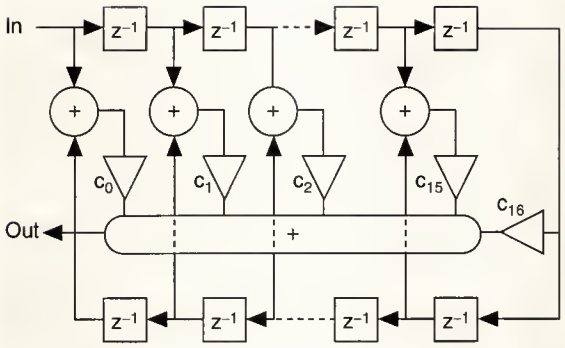
The internal structure of an execution unit, which is based on the generic template shown in Figure 2, remains independent of its function. The operation unit is the functional element of an execution unit. The available operation units[31,32] can be divided into

- arithmetic units: arithmetic and logic (ALU), multiplier-accumulator, programmable logic, and address calculation; and
- memory units: RAM and ROM.

An operation unit receives input operands from a pair of local register files, which connect to a number of buses through bus multiplexers. The presence of local register files reduces bus communication and potential bus transfer conflicts. In each machine cycle every operation unit can produce a result, which transfers to a bus through a tristate bus driver. A number of local instruction registers, which are directly connected to the output of the micro-ROM, control the internal execution unit elements. Some of the operation units (ALU, address calculation, and programmable logic) generate status flags that the control unit evaluates as branch conditions.

The reconfigurable and parameterized Piramid target architecture model facilitates various optimizations during architecture synthesis.

Architecture synthesis. The generation of a Piramid architecture instance is an optimization process that starts from a functional specification in Silage.[33] This functional language specifies what the application is without specifying how it must be implemented in hardware. Figure 3a shows the signal flow graph of a symmetrical transversal filter. As can be seen in the example, the Silage description is a one-to-one reflection of the signal flow graph. The Silage description



(a)

```
#define w16 num<16, 15>
#define w17 num<17, 15>
#define w22 num<22, 15>
#define car num<12, 11>[16]

func main(in: w16) out: w16 =
w16: x;
w17[16]: s;
w22[17]: t, y;
car: c;
begin
#include "c.def"
   x = in;
   y[0] = w22(0);
   (i : 0...15) : :
   begin
      s[i] = w17(x@i+x@(32–i));
      t[i] = w22(c[i]*s[i]);
      y[i+1] = w22(t[i]+y[i]);
   end;
   t[16] = w22(c(16)*x@16);
   out = w16(t[16]+y[16]);
end;
```

(b)

Figure 3. Symmetrical 32-tap FIR filter: signal flow graph (a) and Silage description (b).

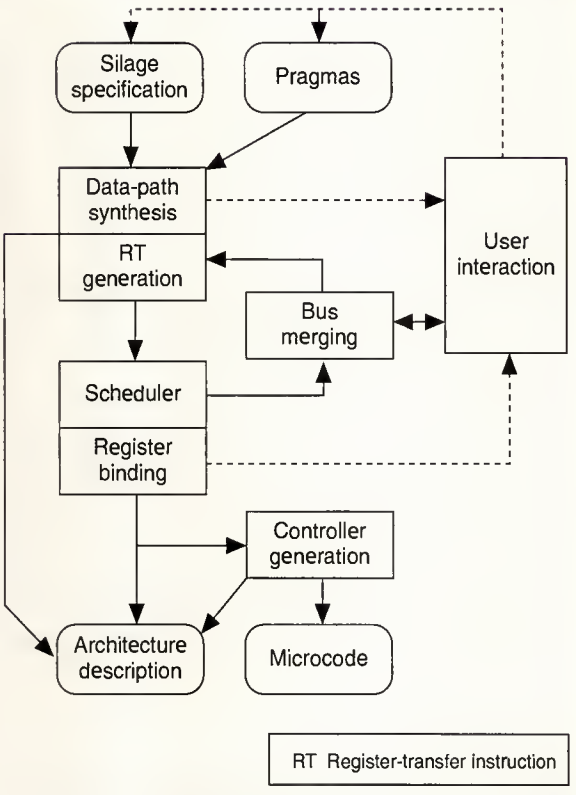


Figure 4. Piramid architecture synthesis.

defines a signal as an infinite sequence of samples. A Silage specification consists of a set of signal equations, rather than assignments, which constitute a time-invariant relation between input and output signals. A unique equation defines each signal, so the order of equations is irrelevant.

An essential feature of Silage is strong typing. The type w16 defined by num<16,15> in Figure 3b specifies a signal represented by a 16-bit two's-complement number with 15 bits after the binary point. Silage permits loop iterations in which signals and constants of array type are frequently used. Functions can be declared in Silage, though this just shortens the specification because the compiler does not preserve function hierarchy. An important construct in Silage is based on the @ operator, which specifies sample-delay functions. The signal operation $s@n$, for instance, specifies a signal s delayed over n sample delays. The functional aspect of the Silage language is of crucial importance since it gives the Piramid architecture synthesis tools maximal freedom. The designer can effectively exploit this freedom by interacting with the compiler (described later).

Figure 4 depicts the design flow of Piramid architecture synthesis. It starts from a Silage specification that is validated by Silage time-domain simulation and optional frequency-domain analysis. Then, generation of a DSP architecture and microcode takes place in subsequent steps as follows:

1) Rule-based data path compilation determines the required execution unit types. By default, Piramid constructs a minimal data path since hardware cost is the primary optimization constraint. A fully connected bus network that is defined to avoid bus communication conflicts assures that every execution unit has write access to a private data bus and can read operands from any data bus.

2) All required operations translate into a set of register-transfer instructions, which directly relate to the generated data path. These instructions reflect operations explicitly visible in Silage, as well as instructions dealing with the transfer of internal variables (address generation, operand fetches) or with loop handling.
3) The scheduler orders these instructions in time in such a way that the total number of required clock cycles is as small as possible.
4) Since the fully connected bus network is redundant in many cases, bus merging is available to combine buses that often are not used simultaneously. Bus merging involves rescheduling of all instructions.
5) Register variables are assigned to fields in register files based on a lifetime analysis. The dimensions of all register files are known after this register-binding step.
6) This final step generates the architecture in detail. It expands execution units down to the level of primitive building blocks and generates the controller architecture and bit-level bus network. It translates the bit-level microcode from the set of register-transfer instructions, including condition handling and the generation of branch addresses.

In principle these steps will generate a complete architecture and associated microcode. However, in most cases the designer wants to control some of the synthesis steps to change the trade-off between hardware cost and speed performance. Mainly, this trade-off depends on a comparison between the cycle budget and the cycle count for a microprogram iteration. The ratio of the feasible clock frequency and the sampling rate determines the cycle budget. The bus merger reduces the number of buses at the cost of additional cycles if the cycle count is smaller than the cycle budget. If the cycle count exceeds the cycle budget, the designer can use an interactive pragma mechanism to increase the DSP's processing power by forcing an extension or a modification of the data path. Pragmas are user directives to the compiler that the designer uses to improve synthesis efficiency in terms of processing power or hardware cost. The set of pragma commands available to the designer includes among others

- *Allocation*, which forces the selection of one or more additional execution units. It can be used to include a new execution unit type or an extra execution unit of an already selected type.
- *Assignment*, which forces operations to be performed by specific execution units.
- *In-place*, which forces variables with nonoverlapping lifetimes to share the same RAM locations.

The pragma mechanism and the bus merger offer a first level of user control that facilitates trade-offs between processor cost and performance. To achieve results with improved efficiency, an architecture designer may want to control some of the synthesis steps to fine-tune the architecture to an application. This level of user interaction goes beyond the use of pragmas. Detailed user control could have been implemented by adding a user interface to every architecture synthesis tool. However, our adopted approach controls architecture synthesis by observing the implications of manipulations in the Silage description. This approach is based on the strategy, "what you write is what you get." The resulting advantages are a simpler compiler, easy control, and more efficient results. We list some examples of manipulations in the Silage description.

- *It can be profitable to exploit the structural properties of a signal-processing algorithm.* The example presented in Figure 3 shows that we can use the property of symmetry to reduce the number of multiplications at the cost of extra additions.
- *In many applications, the algorithm can be easily split into a number of more-or-less independent parts.* It may be advantageous to preserve this hierarchy during scheduling as it normally reduces lifetimes of internal variables that are only stored during the relatively short execution of subalgorithms. The hierarchy of an algorithm can be indicated in the Silage description by enclosing operation clusters using Begin and End statements. Several benchmarks have shown that this approach significantly reduces register cost and cycle count. A similar aspect of signal-flow manipulation relates to condition handling, which often dominates controller cost.
- *The initial Silage specification may contain a variety of signal types, which will all be preserved during architecture synthesis.* This approach may lead to a substantial bus interface overhead. In most cases, however, signal types associated with a certain execution unit instance can be made equal without a loss of performance of the original specification.
- *Sometimes it may be profitable to force a certain variable into a specific execution unit.* In an ALU, for instance, a variable can be added to a constant 0. The compiler does not overrule these "dummy" operations on purpose. Consequently, we can limit the size of the bus network at the cost of some machine cycles. This approach complements bus merging, because the bus merger does not modify the routing of internal variables through execution units.
- *A large number of intermediate variables are stored in register files for a number of clock cycles determined by their respective lifetimes.* Register files are a significant cost factor, since this type of foreground memory is an order of magnitude more expensive than background

RAM. Consequently, it is often profitable to read an internal variable from RAM several times. This multiple fetch must be explicitly specified in Silage.

A Piramid operation example appears in the adjacent box.

The optimization strategies built into the compiler include the use of pragmas and the listed Silage manipulations. However, some applications give poor results in chip area or cycle count due to the general-purpose character of the available set of execution units. Although the parameters of these units afford considerable flexibility, they still suffer the following fundamental limitations:

- In an execution unit only a limited number of operations can be performed within a single clock cycle, even when ample time exists for the operations to execute within a single cycle.
- Some specific operations cannot be mapped efficiently onto an execution unit, for example, complex bit-level operations.[32]
- Some situations still exist in which execution unit instances display hardware overhead. For example, an ALU always contains a carry-generation unit even when only logic operations are assigned to it.

These limitations inspired the introduction of application-specific execution units (AXUs) in Piramid. The designer must identify operations or clusters of operations that are executed repetitively and do not map efficiently onto the available set of execution units. Such operations are isolated in the Silage

Piramid example—CD error corrector

The design of an error corrector for a CD digital audio system illustrates the operation of Piramid. The error corrector is a cross-interleaved Reed-Solomon decoder (CIRC).[34-35] Its functional block diagram (Figure A) contains two Reed-Solomon decoders (C1, C2) and a number of delay functions. The algorithm repeats on a frame basis. An input frame consists of 24 data and eight parity symbols of 8 bits each. C1 and C2 can detect and correct erroneous symbols, though the CD code distance and the adopted error correction strategy limit the number of correctable errors.

Error location and correction are based on the calculation of four Syndrome values per frame;[36] frames are "deinterleaved" between C1 and C2 to distribute burst errors over different frames before they enter C2. At the output of the "descrambling" function, the error corrector generates six 16-bit stereo audio samples per frame, and an error flag every sample to indicate whether the sample is correct.

Frames repeat at a rate of 7.35 kHz. At a target clock frequency of 10 MHz the cycle budget amounts to 1,360 cycles. The Piramid architecture synthe-

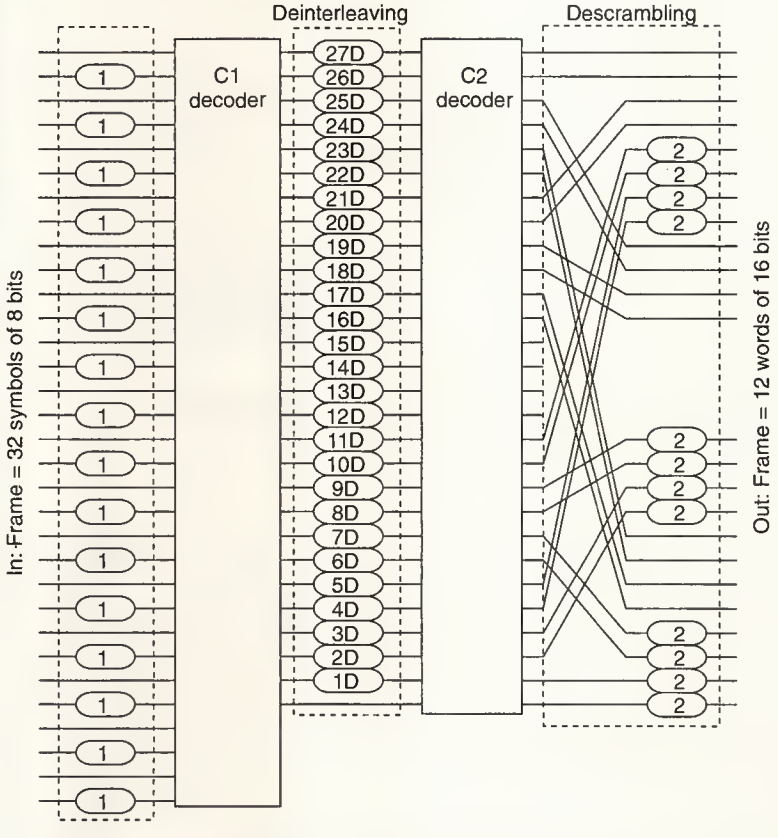


Figure A. Error corrector functional block diagram.

Piramid example (continued)

sis tools generate the data path shown in Figure B. We selected some execution unit types twice to meet the cycle budget constraint. The resulting cycle count is 1,029.

The programmable logic units implement all the C1 and C2 arithmetic operations. Depending on the Syndrome values, they generate status flags, which are used as branch conditions to control the error correction procedure. The address calculation unit and RAM 1, which measures 16 Kbits, carries out the deinterleaving function. The address calculation unit also checks loop bounds. Looping in the microcode limits the length of micro-ROM to just 145 instruction words.

The layout of the 1-micron CMOS error corrector shown in Figure C measures 27 mm^2 and contains 140,000 transistors. This example shows that Piramid architecture synthesis efficiently supports block processing, complex bit-level and word-level arithmetic, delay functions, and complex decision-making functions.

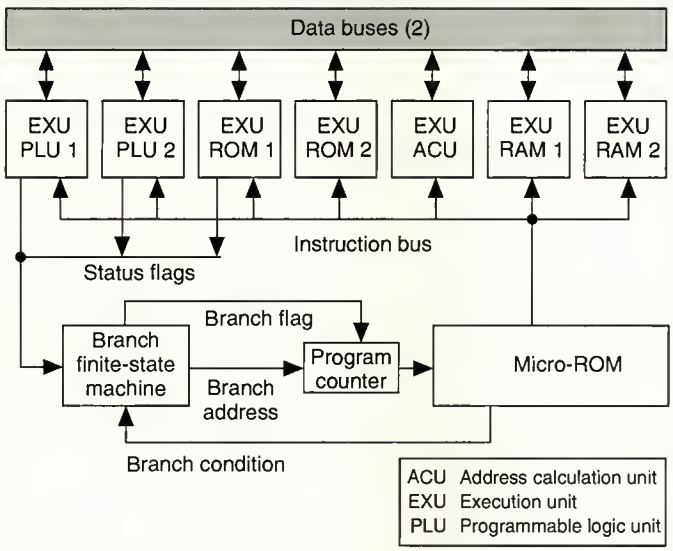


Figure B. Piramid-generated error corrector architecture.

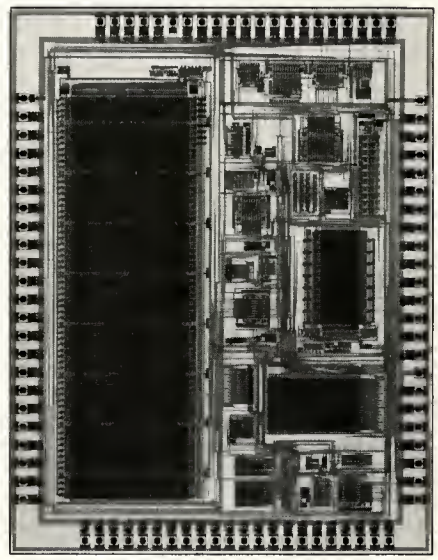


Figure C. Error corrector mask layout.

description. For this purpose Piramid assigns function declarations to AXUs by means of pragmas. Because it ignores pragma statements, the Silage simulator still validates the complete behavior. The architecture synthesis tools interpret AXU function declarations as macroinstructions and treat them as integral execution unit operations. Consequently, AXUs are treated similarly to regular execution units during data path compilation and scheduling. At the architecture level the designer describes the internal structure of an AXU. Then the AXU is imported as an operation unit within a regular execution unit template with local register files and bus interface circuitry. Various design examples using Piramid and AXUs have illustrated that AXU data path elements significantly reduce chip area and cycle count.

The Phideo compiler

The Phideo compiler[29] supports digital video applications such as HDTV in which sampling rates typically approach the maximum feasible clock rates in state-of-the-art CMOS fabrication processes. To meet the associated high throughput requirements, designers based Phideo on a target architecture composed of a number of application-specific, hardwired processing units, distributed memory units with associated address-generation units, and distributed control.

The sampling frequencies associated with digital video applications are much higher than those discussed earlier. In many cases they are comparable to clock frequencies employed in CMOS circuits, and in some cases they are even higher. In addition, various sampling frequencies often occur

in the same application. The standard 13.5-MHz frequency for the luminance signal Y in television applications is a typical sampling rate, though some applications require even higher sampling frequencies. A line-doubling application, for instance, requires twice the standard frequency, that is, 27 MHz. HDTV, with twice the horizontal resolution of ordinary television signals and twice as many lines and fields, requires a sampling frequency of 8×13.5 MHz, or 108 MHz. Lower sampling frequencies also can occur. For example, for each of the chrominance signals (U,V) one fourth of the standard luminance frequency of 3.375 MHz is used.

In digital video processing applications, memories dominate the chip area.[37] Processing at the rate the samples enter the chip achieves the essential minimization of memory cost. Indeed, even a small mismatch between the sampling rate and the processing rate can lead to an explosion in memory costs.

A closer look at video applications reveals two different types of functions. The first type do not change the data values of the samples and relate only to temporal behavior, such as pixel delays, line and field delays, and rate changes. The second type are processing functions that change the value of the samples. They include all arithmetic and logic functions.

Target architecture. The characteristics of video applications led us to develop the Phideo target architecture, which is shown in Figure 5. We must introduce a high degree of parallelism into the architecture to obtain the throughput requirements. Hence, Phideo uses distributed arithmetic and memory units that make the architecture essentially different from the Piramid architecture. Each chip designed with Phideo fits into this general architecture concept. There may, however, be considerable differences, such as the number of processing units or numbers and types of memories and address generators.

The Phideo target architecture directly reflects these two types of video functions: There is a separation between processing and timing functions. Processing units implement the former, and memory units, the latter. Processing units implement all arithmetic and logic operations as well as data-dependent operations; they can also realize small pixel delays. Consequently, processing units can be quite complex. Memory units have a global aspect; that is, different processing units can have read and write access to the same memory location in a memory unit. This aspect differs fundamentally from the Piramid architecture. Dedicated address generator units generate addresses for the memory units, and multiplexed connection networks support communication.

Architecture synthesis. The Phideo compiler generates an implementation for a given functional specification according to the strategy expressed in the design flow shown in Figure 6. The first step clusters operations. The result of this partitioning is a definition of parts of the algorithm, which have to be mapped onto the same type of processing unit.

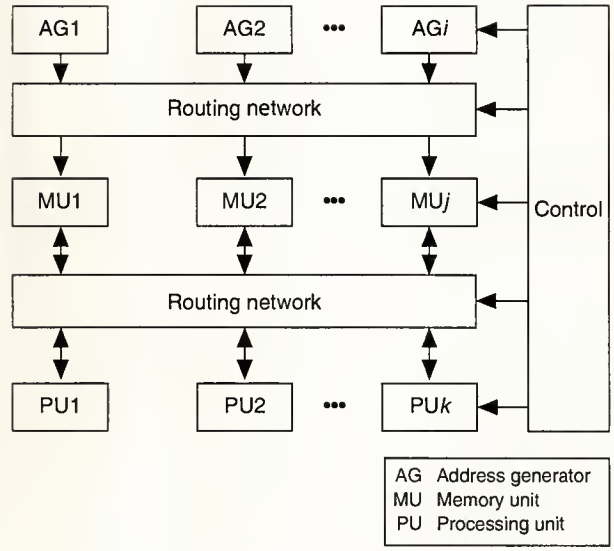


Figure 5. Phideo architecture model.

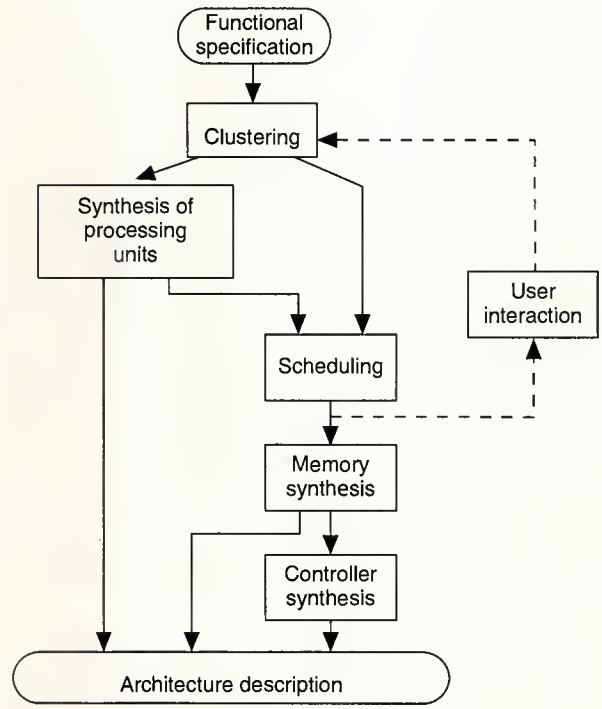


Figure 6. Phideo architecture synthesis.

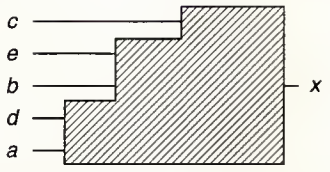


Figure 7. Processing unit example.

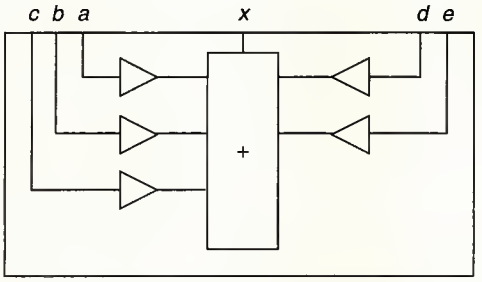


Figure 8. Time shape example.

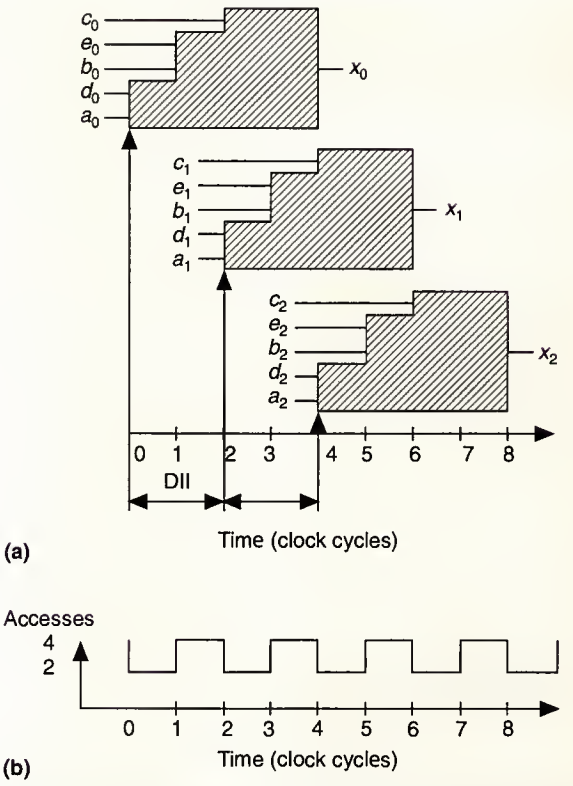


Figure 9. Time shapes (a) and access profile (b).

Clustering is an interactive step in which the designer plays an important role, as discussed later. In this phase the designer also specifies the clock frequency based on sampling rate considerations.

The second step synthesizes the processing units. To obtain the required throughput, designers must optimize the performance of the processing units with a retiming technique.[38,39] This tool involves positioning registers throughout the data path in such a way that delay paths balance and meet the specified clock frequency. Additional registers may be shifted into the network from the input or output ports. These extra registers only modify the temporal behavior of the processing units. The time shape—which shows the clock-cycle skew between a processing unit's inputs and outputs relative to a reference point called its start time—illustrates this technique. Figure 7 shows the time shape of a processing unit with five inputs and one output (shown in Figure 8). The skew between the different inputs and outputs is indicated. We must not confuse the skew with the data introduction interval (DII), which is the interval between subsequent processing unit input sampling times.

The next task to be performed by the compiler is scheduling.[40-41] The scheduler defines the clock cycles at which the processing units accept input data. The example illustrated in Figure 9a depicts three time shapes on a time axis, representing three different executions of the same type of processing unit. In this example the data introduction interval equals two clock cycles, and the first execution starts at clock cycle 0. The scheduler determines these numbers according to the following criteria:

- Minimization of the number of PU instances.
- Minimization of the communication bandwidth between processing units and memory units. This is a crucial optimization criterion for video applications; therefore, memory accesses will be spread as homogeneously as possible as a function of time. Figure 9b shows the corresponding memory access profile.
- Minimization of the total amount of storage of intermediate data values.

The Phideo architectural model allows different processing units to have independent start times and different data introduction intervals. For example, one processing unit can start every clock cycle, while a second unit starts every even cycle and a third starts every odd cycle. The designer evaluates the occupation ratios of the processing units. To improve these ratios, the designer can iterate on the clustering of operations:

- Different functions can be mapped onto the same processing unit. Additional control signals required to select the correct function at the proper time may introduce conflicts and necessitate rescheduling.
- Hardware resources may be shared within a processing unit with time multiplexing.
- Processing units that do not have to be continuously active can run at a lower repetition rate, leading to a significant reduction of the number of internal pipeline registers introduced by retiming. This step may also reduce power dissipation significantly.

After one or more iterations of operation clustering and scheduling, the variables to be read from or written to memory for each clock cycle are known. Memory synthesis must then ensure that bandwidth requirements are met, a step which involves the generation of several memories in parallel and an appropriate communication network. Furthermore, variables must be carefully assigned to memories. Memory synthesis solves the bandwidth problem using a limited number of memory types such as a single-port SRAM or a dual-port register file. The applied memory synthesis techniques[42] include memory decimation, accelerated read, and delayed write.

The next step binds variables to memory locations, making all memory dimensions known, and the last step in memory synthesis generates the address sequences.[43,44] Finite-state machines implement the address generator units, which have an enable signal to control the state transitions and a reset signal to start the sequence from a well-defined state.

Finally, the (Phideo) compiler generates the controller. Several functions may be implemented in the same processing unit, and the controller is responsible for the selection of the correct function in each clock cycle. The controller also generates the control signals for the multiplexed communication network, the read and write signals for the memory units, and the enable and reset signals for the address generation units. An important characteristic of the controller signals is their repetitive nature, a direct consequence of the fact that the start times for the processing units are equidistant in time. The box illustrates a Phideo operation.

Layout synthesis. The output of the Piramid and the

Phideo example—scan conversion for a 100-Hz TV

We can illustrate the operation of Phideo with a conversion algorithm from an interlaced scan format to progressive scan format.[45] The schematic in Figure D presents this conversion, which produces an improved picture quality. Because the number of lines per screen is doubled, it puts out samples at 54 MHz, twice the input rate (27 MHz), and an internal rate conversion takes place. The computation of the additional lines depends on the local direction in the original picture. We assume that the processor runs at 27 MHz and produces two output samples (even/odd) each clock cycle.

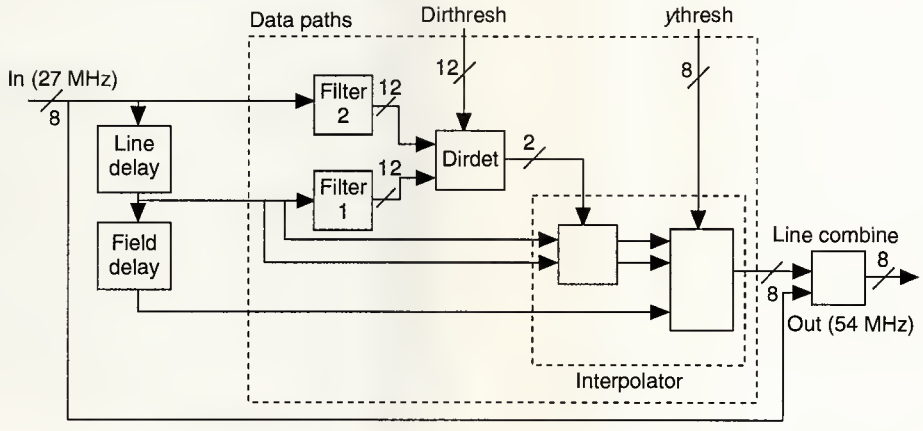


Figure D. Progressive scan functional block diagram.

continued on next page

Phideo example (continued)

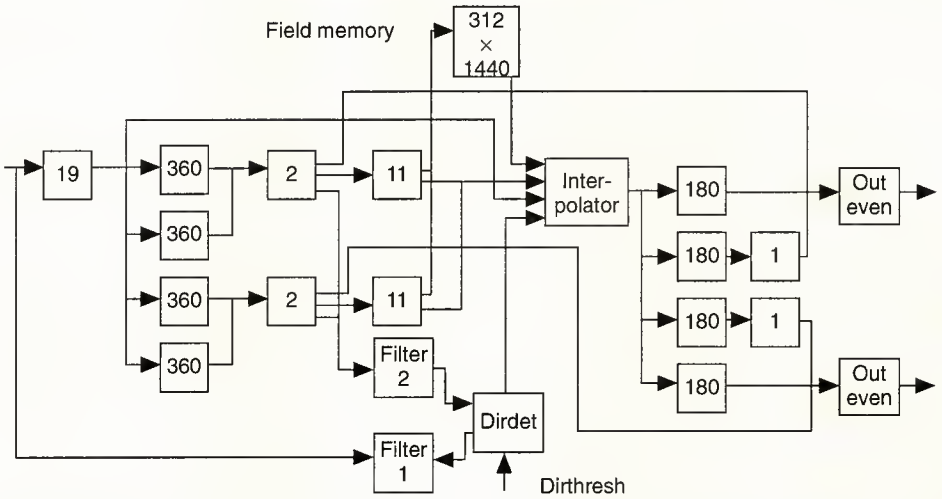


Figure E. Phideo-generated progressive scan architecture.

Figure D identifies three functions: a nine-tap FIR filter, a direction-detection function, and an interpolation function. The main on-chip delay functions are line delay and line combine. We consider the field delay as an off-chip block with extra inputs from and outputs to the chip.

A mapping of the conversion algorithm onto Philips VSP ICs requires 16 chips, but Phideo can map the algorithm onto one chip, retiming the three processing units to reach the clock frequency of 27 MHz. Figure E shows the detailed structure of the processor after scheduling and memory synthesis. Here, the compiler generates a distributed network of 15 memory units and indicates the minimum number of words in a memory, determined by lifetimes of the samples to be delayed, in each case. Four 360-word RAMs implement the line delay and the time compression of the input signal. This combination of timing functions saves half of the line memory, that is, 720 words. Four 180-word RAMs perform time compression on the processed signal. Some other small delays compensate for delays added during retiming of the data paths. All other memories implement small delays over a few clock cycles to solve access conflicts in the preceding RAMs. For this application we used counter addressing for all memory units.

The processing units take up 11 mm^2 on the chip, the memory units 16 mm^2, and the address generation units 3 mm^2. The chip area including routing measures 38 mm^2 in a 1-micron CMOS process. Except for the memories, we used a standard-cell layout style (Figure F).

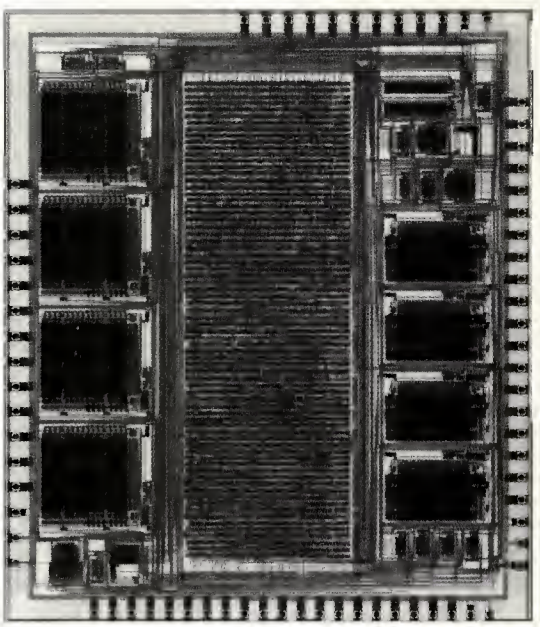


Figure F. Progressive scan mask layout.

Phideo compilers is a bit-true description of an application-specific architecture; see Figure 10. Both compilers employ a common layout synthesis environment for the translation of an architecture description to layout. The layout synthesis environment facilitates rapid generation of area and performance figures. Designers can use these figures to estimate cost and performance at an early stage during architecture synthesis. The first step in layout synthesis is a decomposition of the architecture into primitive bit-level blocks. For this purpose the layout design system uses hierarchical expansion, logic synthesis, and retiming tools.[46]

The result is a so-called netlist comprising all required library elements and their connections. The second step in layout synthesis generates layout patterns for all modules in the netlist. This step distinguishes two kinds of modules.

- Macro blocks are assembled from a library of parameterized module generators.[47] Examples are RAMs and ALUs.
- Cell-based blocks are assembled by conventional standard-cell placement and routing techniques.

Finally, the layout design system generates a floor plan[48] and assembles all layout patterns into a complete mask layout, including all connections, supplies, clocks, and test routing. (SCII from Silvar Lisco in California assembles the layout.)

The completely automatic layout synthesis shortens design time and relieves the designer of a tedious and error-prone task. The same argument holds for design verification. In our opinion it is impossible to guarantee correctness by constructing designs to accommodate every possible synthesized architecture.

Consequently, our process makes verification tools—including conventional simulators at register, gate, and switch levels—available at all stages in the synthesis process. Since it is principally impossible to detect all possible design errors by simulation, the layout synthesis uses a supplementary verification technique based on hierarchical structure abstraction.[49] This technique reconstructs the original architecture from a transistor-level netlist extracted from layout. In addition, accurate timing verification is available.

The generation of a complete layout, starting from an architecture description, takes a few hours on a normal workstation; its verification is a matter of days. Iterations on architecture and layout synthesis for efficiency improvements take a few weeks.

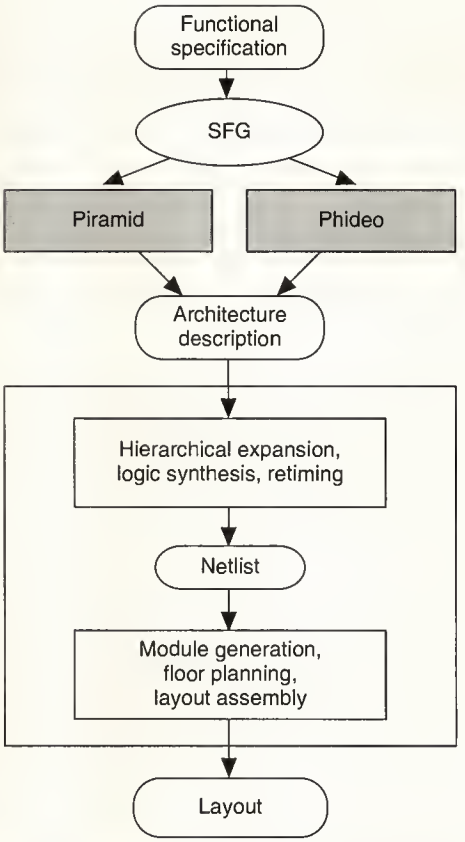


Figure 10. Complete design system (Piramid/Phideo/back-end).

THE LIFE CYCLE PHASE OF THE ENVISAGED APPLICATION strongly determines the decision to employ a general-purpose or a more application-specific approach to realizing a digital signal processing system. Relevant decision parameters include the expected production volume, the acceptable development overhead, and the required application flexibility.

Stringent time-to-market requirements, shorter product lifetimes, and lower profit margins necessitate highly effective design approaches. High-level synthesis techniques facilitate accelerated design space exploration and optimization of a chosen design. Strategic development of these techniques should reduce development time and costs.

The literature has already demonstrated the effectiveness and industrial relevancy of architecture compilers like Piramid and Phideo.[50-52] These compilers offer a variable degree of user interaction, which allows designers to seek an optimal balance between development time and fabrication costs.

Acknowledgments

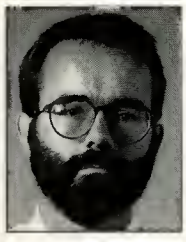
The high-level synthesis results we described testify to the commitment of Philips Research scientists cooperating to reduce IC development overhead. These results are partly based on the ESPRIT projects 97 and 2260 (Sprite). In the former project IMEC, Leuven, Belgium, and Philips Research Labs

closely cooperated to develop a first version of the Cathedral-II compiler. Philips Research Labs used this version as the starting point in developing the Piramid Design System. As a contribution to the latter project, Philips Research Labs developed the Phideo compiler.

References

1. Special Issue on the Compact Disc Digital Audio System, *Philips Technical Review*, Vol. 40, No.9, 1982.
2. H. Sahm et al., "RISS-16, Realization of a DSP Optimized for Digital Mobile Radio Systems," *Proc. Euro-ASIC*, CS Press, Los Alamitos, Calif., 1991, pp. 36-40.
3. W.P.G. Crooijmans, "From Algorithm to VLSI for HDTV," *IEEE Trans. Consumer Electronics*, Vol.37, No.4, Nov. 1991.
4. G.C.P. Lokhoff, "DCC—Digital Compact Cassette," *IEEE Trans. Consumer Electronics*, Vol. 37, No. 3, Aug. 1991, pp. 702-706.
5. M. Alard and R. Lassalle, "Principles of Modulation and Channel Coding for Digital Broadcasting for Mobile Receivers," *European Broadcasting Union Review*, No. 224, Aug. 1987, pp. 3-25.
6. E.A. Lee, "Programmable DSP Architectures," Parts I and II, *IEEE ASSP*, Vol.5, No. 4, Oct.1988, and Vol. 6, No.1, Jan.1989, pp. 4-19 and pp.4-11.
7. E.A. Lee, "Programmable DSPs: A Brief Overview," *IEEE Micro*, Oct. 1990, Vol. 10, No. 5, pp.14-16.
8. J.L. van Meerbergen et al., "A 2-mm CMOS 8-MIPS Digital Signal Processor with Parallel Processing Capability," *Proc. Int'l Solid-State Circuits Conf.*, Lew Wenner, Coral Gables, FL 33134, 1986, pp. 84-85.
9. A.H.M. van Roermund et al., "A General-Purpose Programmable Video Signal Processor," *IEEE Trans. Consumer Electronics*, Aug. 1989, pp. 249-257.
10. R. Woudsma et al., "VHDL Modelling of Reconfigurable DSP Architectures," *Proc. IEEE Benelux & ProRISC Workshop on Circuits, Systems and Signal Processing*, STW Technology Foundation, Utrecht, The Netherlands, Apr. 1992, pp. 131-138.
11. L. Matterne et al., "A Flexible High-Performance 2D Discrete Cosine Transform IC," *Proc. Int'l Symp. Circuits and Systems*, Vol. 2, IEEE, New York, 1989, pp. 618-621.
12. P. Marwedel, "A New Synthesis Algorithm for the MIMOLA Software System," *Proc. 23rd Design Automation Conf.*, ACM/IEEE, New York, June 1986, pp. 271-277.
13. P.G. Paulin, J.P. Knight, and E.F. Girczyc, "HAL: A Multi-Paradigm Approach to Automatic Datapath Synthesis," *Proc. 23rd Design Automation Conf.*, June 1986, pp. 263-270.
14. A.C. Parker, J. Pizarro and M. Mlinar, "MAHA: A Program for Datapath Synthesis," *Proc. 23rd Design Automation Conf.*, June 1986, pp. 461-466.
15. F.D. Brewer and D.D. Gajski, "Knowledge-Based Control in Micro-Architecture Design," *Proc. 24th Design Automation Conf.*, June 1987, pp. 203-209.
16. H. Trickey, "Flamel: A High-Level Hardware Compiler," *IEEE Trans. CAD*, Vol. CAD-6, No. 2, Mar. 1987, pp. 259-269.
17. R.K. Brayton et al., "The Yorktown Silicon Compiler," in *Silicon Compilation*, D.D. Gajski, ed., Addison Wesley, Reading, Mass., 1988, pp. 204-311.
18. L. Stok and R. van den Born, "EASY: Multiprocessor Architecture Optimization," *Proc. Int'l Workshop on Logic and Architecture Synthesis for Silicon Compilers*, G. Saucier and P.M. McLellan, eds., Elsevier Science Publishers, Amsterdam, 1988, pp. 313-328.
19. T. Fuhrman et al., "The System Architect's Workbench," *Proc. Int'l Symp. Circuits and Systems*, Vol. 4, 1990, pp. 2584-2587.
20. R. Camposano and W. Rosenstiel, "Synthesizing Circuits from Behavioural Descriptions," *IEEE Trans. CAD*, Vol. 8, No. 2, 1989, pp. 171-180.
21. P. Duzy et al., "CALLAS—Conversion of Algorithms to Library Adaptable Structures," *Proc. VLSI*, IFIPS, New York, 1989, pp. 197-208.
22. P. Denyer and D. Renshaw, *VLSI Processing: A Bit Serial Approach*, Addison Wesley, 1985.
23. C.B. Shung et al., "An Integrated CAD System for Algorithm-Specific IC Design," *IEEE Trans. CAD*, Vol. 10, No. 4, Apr. 1991, pp. 447-482.
24. B.S. Haroun and M.I. Elmasry, "Automatic Synthesis of a Multibus Architecture for DSP," *Proc. IEEE Conf. Computer-Aided Design*, CS Press, Nov. 1988, pp. 44-47.
25. H. De Man et al., "Architecture-Driven Synthesis Techniques for VLSI Implementation of DSP Algorithms," *Proc. IEEE*, Feb. 1990, pp. 319-335.
26. J. Van Meerbergen and H. De Man, "A True Silicon Compiler for the Design of Complex ICs for Digital Signal Processing," *Philips Technical Review*, Vol. 44, No. 7, 1989, pp. 218-230.
27. R. Woudsma et al., "Piramid: An Architecture-Driven Silicon Compiler for Complex DSP Applications," *Proc. Int'l Symp. Circuits and Systems*, Vol. 4, 1990, pp. 2596-2600.
28. J.Rabaey and M. Potkonjak, "Resource Driven Synthesis in the HYPER System," *Proc. Int'l Symp. Circuits and Systems*, Vol. 4, 1990, pp. 2592-2595.
29. P.E.R. Lippens et al., "Phideo: A Silicon Compiler for High-Speed Algorithms," *Proc. European Conf. Design Automation*, CS Press, 1991, pp. 436-441.
30. J.L. van Meerbergen, "Efficient Controller Architectures for DSP Compilers," *Proc. Custom Integrated Circuits Conf.*, IEEE, New York, 1988, pp. 14.5.1-14.5.4.
31. J. Huisken et al., "Design of DSP Systems on Silicon Using the Piramid Library and Design Tools," in *VLSI Signal Processing III*, R.W. Brodersen and H.S. Moscovitz, eds., IEEE, New York, 1988, pp. 319-329.
32. P. de Bakker, A. Delaruelle, and B. De Loore, "Design of a Programmable Logic Unit, a New Block for Signal Processing," *Proc. European Solid-State Circuits Conf.*, Sep. 1988, pp. 58-61.

33. P. Hilfinger, "A High-Level Language and Silicon Compiler for Digital Signal Processing," *Proc. Custom Integrated Circuits Conf.*, IEEE, New York, May, 1985, pp. 213-216.
34. A.M. Michelson and A.H. Levesque, *Error-Control Techniques for Digital Communication*, Wiley-Interscience, New York, 1985.
35. L.B. Vries and K. Odaka, "CIRC—the Error Correcting Code for the CD Digital Audio System," *Digital Audio Collected Papers*, 1982.
36. A. Delaruelle, "The Design of a Syndrome Generator Chip Using the Piramid Design System," *Proc. European Solid-State Circuits*, Sep. 1988, pp. 256-259.
37. K. Huizer et al., "A Programmable 1400-MOPS Video Signal Processor," *Proc. Custom Integrated Circuits Conf.*, IEEE, New York, May 1989, pp. 24.3.1-24.3.4.
38. C.E. Leiserson, F.M. Rose, and J.B. Saxe, "Optimizing Synchronous Circuitry by Retiming," *Proc. Third Caltech Conf. Very Large Scale Integration*, Computer Science Press, Rockville, Md., 1983, pp. 87-116.
39. A. van der Werf et al., "Hierarchical Retiming Including Pipelining," *Proc. Int'l Conf. VLSI*, Elsevier Science Publishers, Aug. 1991, pp. 11.2.1-11.2.10.
40. W. Verhaegh et al., "Improved Force-Directed Scheduling," *Proc. European Design Automation Conf.*, CS Press, Feb, 1991, pp. 430-435.
41. G. Essink et al., "Scheduling in Programmable Video Signal Processors," *Proc. IEEE Conf. Computer-Aided Design*, CS Press, Nov. 1991, pp. 284-287.
42. P. Lippens et al., "Memory Synthesis for High-Speed DSP Applications," *Proc. Custom Integrated Circuits Conf.*, IEEE, New York, May 1991, pp. 11.7.1-11.7.4.
43. D.M. Grant, P.B. Denyer, and I. Finlay, "Synthesis of Address Generators," *Proc. IEEE Conf. Computer-Aided Design*, CS Press, Nov. 1989, pp. 116-119.
44. D.M. Grant and P.B. Denyer, "Address Generation for Array Access Based on Modulus m Counters," *Proc. European Conf. Design Automation*, CS Press, 1991, pp. 118-122.
45. T. Doyle and M. Looymans, "Progressive Scan Conversion Using Edge Information," *Proc. Third Int'l Workshop on HDTV*, 1989.
46. A. van der Werf et al., "Flexible Datapath Compilation for Phideo," *Proc. Euro ASIC*, CS Press, May 1991, pp. 178-183.
47. H.A.V. Janssen et al., "GrapMG: Cost-Effective Module Generation," *Proc. European Solid-State Circuits Conf.*, 1989, pp. 68-71.
48. F.M.J. de Bont et al., "Placement of Shapable Blocks," *Philips J. Research*, No. 43, 1988, pp. 1-19.
49. T. Kostelijk and D. De Loore, "Automatic Verification of Library-Based IC Designs," *IEEE J. Solid-State Circuits*, Vol. 26, No. 3, Mar. 1991, pp. 394-403.
50. R. Woudsma and A. Delaruelle, "The Design of DSP Components for the CD Digital Audio System Using Silicon Compilation Techniques," *Proc. Custom-Integrated Circuits Conf.*, 1989, pp. 20.4.1-20.4.4.
51. D. Genin, "The Mentor Graphics/EDC DSP Station," *Proc. Benelux Symp. State-of-the-Art Digital Signal Processing*, Mar. 1991.
52. E. Roza, "On the Application of Architectural Synthesis in the Design of High-Volume Production ICs," *Proc. Synthesis and Simulation Meeting and Interchange*, SASIMI Workshop, Apr. 1992, pp. 313-323.



Rob Woudsma is a senior research scientist in the Digital VLSI Department at Philips Research Labs in Eindhoven, The Netherlands. His research activities include MOS circuit design, hardware description languages (such as VHDL), DSP applications and architectures, VLSI design methodologies in general, and high-level synthesis in particular.

Woudsma received an MS degree in electrical engineering and currently is pursuing a PhD degree at Twente University, Enschede, The Netherlands.

Jef L. van Meerbergen, also a senior research scientist in the same department, is presently engaged in architecture synthesis for digital video applications. He has been involved in the design of digital MOS circuits, general-purpose DSPs, and high-level synthesis techniques. He received MS and PhD degrees in electrical engineering from the University of Leuven, Belgium.

Direct questions concerning this article to Rob Woudsma, Philips Research Laboratories, WAY 4-21, PO Box 80000, 5600 JA Eindhoven, The Netherlands; or via e-mail at woudsma@prl.philips.nl.

Analog-Digital Technologies for Mixed-Signal Processing:

The Driving Force to Success for the European Industry

Mixed analog-digital products complement digital processors in a wide range of systems. We describe the development activities at European semiconductor companies and research institutes, both on A/D processes and design tools. The processes include mixed A/D CMOS, high-voltage interface BiCMOS, low-voltage A/D BiCMOS, high-voltage interface BiCMOS, and low-voltage A/D BiCMOS. Tools include analog circuit synthesis and library cells.

Edgard Laes

Herman J. Casier

Eric Schutz

Mietec Alcatel

In today's microelectronics field, embedded processors and digital signal processors—most often designed with CMOS technology—provide plenty of computational power and digital control. However, many applications need an interface processor between the digital processor and the outside world. Obvious examples are analog-to-digital converters, filters, and amplifiers. Combining the digital logic or processor with interface functions provides more integrated and more reliable solutions that need less board space.

We describe some mixed technologies that realize the standard CMOS devices for the logic part and add other devices necessary for analog circuits or high-voltage output devices. These technologies are mixed A/D CMOS, interface BiCMOS (bipolar CMOS), and A/D BiCMOS. To fully exploit the opportunities of mixed technologies, designers must develop design tools so advanced analog cells can be designed and simulated. The work on analog circuit synthesis tools, both at CAD companies and several universities, drastically improves the productivity of mixed-signal circuit designs.

Large European companies are front-edge players in the world market for consumer electronics, telecommunications, and automotive electronics. The integrated circuit needs from these companies keep Europe in the forefront for mixed technologies. We show the link between future product needs and mixed-technologies research and development activities.

A/D CMOS technology

The requirements of the A/D CMOS technologies can be compared to the state-of-the-art processes and to the new process under development. Of interest are cell development activities related to particular application fields.

Requirements. Circuits such as switched capacitor filters, RC filters, and A/D or D/A converters display high accuracy and a large dynamic range. They also require area-efficient precision resistors and capacitors with good matching properties.

In a standard 1-µm CMOS process, the low polysilicon gate-layer resistance of 30 ohms per square severely limits its use as a precision resistor. The low polysilicon-to-metal capacitance about equals the parasitic capacitance to the substrate. Thus logic CMOS processes can integrate low-performance analog functions in a cost-effective way. High-performance analog circuits, however, need extra layers in the process to form accurate linear capacitors and resistors.

To capitalize on the availability of digital libraries that contain well-defined cells and megacells, designers need a modular process that lets them plug in options. However, the process changes to implement the additional devices must not change the characteristics of the logic devices.

0272-1732/92/0800-0034$03.00 © 1992 IEEE

Another key requirement for analog devices is proper matching. Since only a low absolute accuracy of ±20 percent on resistors, capacitors, and transistor drive currents can be achieved in manufacturing, the design of efficient analog circuits requires good matching between identically designed devices.

Another parameter to be defined and optimized is noise: The circuit's dynamic range and the minimum signal that can be resolved in filters or converters are limited by noise.

CMOS A/D development. Mixed A/D CMOS processes in volume manufacturing today have polysilicon gate lengths of 1.5 µm to 2 µm. During the past two years seven European companies worked together under the JESSI program called Joint Logic to advance logic CMOS processes and functional options including analog features. The first target for the mixed process was the analog option for 1-µm CMOS technology. The different partners did not install the same process but worked on similar analog modules and characterization techniques.

Table 1 lists the features of the additional analog devices and specific characteristics relating to analog performance.

Table 1. Devices for analog modules.

Device	Feature	Characteristic
Capacitor	Capacitance/tolerance	0.85 fF/µm² ± 15%
	Linearity from –5V to +5V	< 50 ppm/ V
	Matching for an area larger than 100 µm²	< 0.6%
Hipo resistor	Resistance/tolerance	2,000 ohms/sq ± 20%
	Temperature coefficient	–2,500 ppm/°C
	Matching for L > 100 µm, W > 10 µm	< 0.5%
Low V_t PMOS	Threshold voltage	–0.7V ± 0.15V
	Minimum drawn gate length	1.5 µm
	Current matching in saturation	0.5-2.0% (depends on biasing)
	1/f noise at 100 Hz (10/10 transistor biased at 20 µA)	35 nV/sqrt Hz
NMOS	Threshold voltage	0.7V ± 0.15V
	Current matching in saturation	0.5-2.0% (depends on biasing)
	1/f noise at 100 Hz (10/10 transistor biased at 20 µA)	80 nV/sq rt Hz

Designers realized area-efficient precision resistors with good matching properties in moderately doped polysilicon, further called hipo (high ohmic polysilicon) resistors. By implanting boron in the polysilicon, designers achieved a typical hipo resistance of 2 kohms/square. They further masked the resistor areas in the process from the much stronger phosphorous doping for the gate and the capacitor plates. Interesting features of the hipo resistors are their complete isolation from the substrate, voltage-independent resistance under normal biasing conditions, and negative temperature coefficient. The designers exploited the negative temperature coefficient.

They realized the capacitors in two ways, the first of which used the double-polysilicon capacitor. In this architecture the first polysilicon layer serves as the transistor gate and the capacitor bottom plate. The most convenient choice for the interpolysilicon capacitor dielectric is an oxide grown using the TEOS (tetra ethyl ortho silicate) chemical vapor deposition technique. The advantages of this concept are a low parasitic capacitance, dielectric isolation of both capacitor plates from the substrate, and a low capacitance voltage coefficient of 30 ppm/V. Disadvantages are the worsening of topography and the risk of short formation between adjacent lines in the second polysilicon layer. This last problem is very well known in double-polysilicon processes presently in production. Due to the anisotropic nature of the polysilicon etching process, small wires in the second polysilicon layer could be left over at steps in the first polysilicon layer, resulting in shorts between two adjacent polysilicon-2 lines.

An alternative method realizes the capacitor with a heavily doped N layer in the silicon substrate as the bottom plate and the single polysilicon layer as the top plate. The capacitor dielectric is a thermal oxide grown together with the gate oxide. A disadvantage of this capacitor is the larger and voltage-dependent parasitic capacitance of the bottom plate to the substrate, but the added processing is small and a voltage coefficient below 50 ppm/V can be guaranteed.

The core process PMOS transistor guarantees no leakage at 0V gate biasing over the full temperature range, but low-threshold voltage transistors must be added to increase the maximum voltage swing.

The thermal budget for the analog module's additional process steps is kept small to keep the core logic process device characteristics unchanged.

Device engineers studied the matching of transistors, resis-

Two companies are now working on this 1-μm A/D CMOS process, and a new challenge is appearing at the 0.6/0.5-μm level.

tors, and capacitors extensively. They used a statistical model to relate device-matching parameters such as transistor current to the spread of commonly measured distributions. The current matching for two identical transistors used in saturation in a current mirror is given by:[1]

$$[\text{sigma}(\Delta I)/I]^2 = [\text{sigma}(ß)/ß_M]^2 + 4[\text{sigma}(V_t)/(V_g - V_t)]^2$$

where sigma($\Delta I/I$) is the standard deviation of the relative difference of the transistor pair currents, sigma ($ß/ß_M$) is the mean standard deviation of the transistor pair transconductances, and sigma $V_t/(V_g - V_t)$ is the deviation of the transistor pair threshold divided by the difference of the applied gate voltage with the threshold voltage.

The strength of this model is that it relates current matching to spreads that are controlled in wafer manufacturing. It also models the relation between matching and transistor biasing. Similar work is done for resistors and capacitors. For the 100-μm^2 capacitor the capacitance is less than 0.1 picofarad. Since standard measurement techniques are not accurate enough to measure matching, designers developed a new test structure that translates capacitance matching into voltage matching, which can easily be measured automatically.

Due to the weak avalanche on the drain side of an NMOS transistor, a drain-substrate current must be added to the transistor current. This addition results in a nonlinearity of the drain current in saturation and a decrease in output resistance of the transistor. With drain engineering, designers can keep a minimum electric field at the drain and thus decrease the drain-substrate current component. Transistor models for analog applications must reflect the output conductance, including the above-mentioned weak avalanche drain-substrate currents and the linear-to-saturation transition, without discontinuities in the currents and their derivatives. Also required is accurate modeling of the weak, moderate, and strong inversion regions and the velocity saturation for a wide range of transistor lengths.

The partners brought existing commercial and proprietary models from different companies into research projects for further study and improvements.

Although several design techniques can minimize the impact of noise in the considered signal bandwidth, the 1/f noise remains a fundamental limitation of the MOS transistor. The 1-μm A/D processes contain a lot of plasma processing and mechanical stress in the silicon increases, for instance, resulting from the double-level metal interconnection process. Both effects influence 1/f transistor noise, so designers must also optimize the processes for this aspect. Table 1 lists typical 1/f noise results.

SGS-Thomson Microelectronics and Mietec Alcatel are now working on the 1-μm A/D CMOS process.

Future A/D process development. Digital 0.8-μm and 0.7-μm CMOS processes and cell libraries are now becoming available. The addition of an analog module to these processes creates new opportunities for system integration.

The partners have already begun development of the 0.8-μm analog module; its specification and modular architecture are comparable to the 1-μm process. As just described, characterization and optimization toward matching, linearity, noise, and device stability require a lot of effort before this module can be released to production. We expect production in early 1993.

At the 0.6/0.5-μm level a new challenge is appearing, namely the supply voltage. As already amply demonstrated, the 5V supply for a 0.5-μm transistor does not fulfill the reliability requirements. On the other hand, at 3.0V, the dynamic range for the analog signal seriously decreases. By the end of 1992 we expect the process architecture and transistor characteristics for the 0.6/0.5-μm digital process to be stable. Then we can fix the process for the analog module.

Cell development and analog circuit synthesis. We obtain the mixed A/D process by adding analog modules to a standard digital core process. The digital libraries developed for the standard digital processes are hence also available in these mixed A/D processes. Borel, Moreau, and Samani in this issue (p. 43) extensively describe these libraries, the associated design methodologies, the standardization, and the CAD environment and interfaces, so we will not discuss them further.

The simpler generic digital functions are easily defined and are characterized by a few parameters (intrinsic delay, load dependent delay, power, and so on). This is not true for analog functions in which many parameters are important, such as gain, bandwidth, offset, bias current, power supply rejection ratio (PSRR), common-mode rejection ratio (CMRR), slew rate, settling time, and noise. The trade-off between all these parameters results in a very large number of generic circuits and is exemplified by the enormous variation of specifications described in data books of standard analog components.

Clearly, an analog library of generic functions and macros is not easily defined, and analog circuit synthesis (schematic and layout synthesis) is very important. Unfortunately, ana-

log circuit synthesis is not in such an advanced state as digital circuit synthesis (with the exception of switched capacitor filter synthesis). European companies are very active in this domain, however.

In the past researchers developed extensive standard cell libraries with many different types of generic cells, together with a CAD system for mixed simulation and layout. Even with these large libraries new high-performance cells had to be developed in almost every mixed design. Due to the increased analog performance requirements, more and different design trade-offs must be met, resulting in cells with an increasing performance and an decreasing reusability potential.

As a result, circuit designers developed a smaller library of generic general-purpose cells complemented with a few complex, state-of-the-art analog cells such as references, A/D and D/A converters, and standard signal interfaces. These cells have a high reusability potential, and their development takes too long in the course of a commercial design project. The differential wideband characteristics in Table 2 illustrate the performance that can be reached with 1.0-µm mixed-signal CMOS cells.[2]

Finally, designers now also spend much more effort in analog circuit synthesis. The first part of analog circuit synthesis involves the schematic synthesis. Starting from the specifications, designers optimize circuits and center designs for maximum tolerance against parameter variations. Essential here are, of course, the synthesis software, the analog and/or mixed-mode simulator, and the accurate analog and statistical modeling of the components. Various research programs throughout Europe pursue all these aspects.

During this phase designers rely heavily on the analog and/or mixed-mode simulator and on the accurate analog and statistical modeling of the components. The novice designer uses existing circuit topologies built into the synthesis tool and optimizes them for application, while the expert designer devises new circuit topologies. The novice designer will not be able to reach the ultimate performance that can be reached by the expert designer. The expert on the other hand would rather use selected synthesis tools to speed up the design task and, more importantly, provide more insight into the newly devised circuits.

The layout synthesis begins from the optimized schematic complemented with specific analog layout properties. The challenge here is to simultaneously optimize all the very different analog layout features.

Various European programs offer examples of this approach. They include ADCIS for schematic and layout synthesis, complex cell research, analog technology characterization; AD2000 for state-of-the-art converters and synthesis; AC41 for analog technology modeling, characterization, and reliability; and Plasic for the characterization of plastic packaging effects on analog performance.

Besides these European programs, many other local research programs and commercial or proprietary systems are developed in Europe.[3] Examples are Midas (Philips Consumer Electronics) for analog circuit synthesis of reusable functions, Ariadne (Catholic University of Leuven), and Adam (CSEM) for schematic and layout synthesis, and Mixsim (Mietec Alcatel) and Eldo (Eldo-FAS, Fidel, Eldo-XL from Anacad), and Mozart-Eldo (SGS-Thomson) for mixed-mode simulation.

Table 2. Specification of a differential wideband amplifier.

Characteristic	Value
DC gain	39 dB
Gain bandwidth	800 MHz
Phase margin	53°
Load capacitance	5.0 pF
Input capacitance	1.7 pF
Input offset	± 5 mV
Common-mode output offset	< ± 50 mV
Power consumption	240 mW

Product drivers. The telecommunications industry still drives mixed-signal A/D CMOS technologies not withstanding the digitalization.[4] Today's 1-µm technology supports chips on the line card of public switches, user and subscriber interface chips for N-ISDN and transceiver chips in local area networks. The emerging 0.8-µm and 0.6-µm technologies are necessary for an economical, widespread implementation of digital cellular radio telephony (GSM, the Global System for Mobile, and DECT, the Digital European cordless telephone). They will also help us come to a single-chip, variable bit-rate video codec for high-resolution videophones. For the future high-definition and intermediate-definition television products, submicron mixed-signal CMOS technology provides fast and high-resolution A/D and D/A converters complementing the digital signal processors.[5,6] This technology also supports application for automotive electronics in the field of bus systems and safety control units.

Interface BiCMOS processes

Interfacing between a digital processor and the outside world requires a two-way signal flow and signal translation. Analog signals coming from different sensors and measurement systems must be converted into digital signals for the processor. The control signals returning from the processor need to drive the actuators, which often require high power. Typical applications also contain 100 to 2,500 logic gates that alleviate the need for a separate processor or unloading the processor.

Combining these requirements leads to a high-voltage

Table 3. Typical DBiMOS (Mietec Alcatel) device characteristics.

Device	Characteristic	Value
Bipolar transistor NPN	Beta	100
	BV_{CEO}	> 40V
	BV_{CES}	> 80V
	V_{Early}	250V
	f_T	600 MHZ
Bipolar transistor PNP	Beta	300
	BV_{CEO}	> 40V
	BV_{CES}	> 80V
	V_{Early}	100V
	f_T	7 MHZ
NMOS	Min. gate length	4 µm
	V_t	1.1V
	Transconductance	30 µA/V²
	Linearity of I_{DS} up to V_{DS}	11V
PMOS	Min. gate length	3 µm
	Vt	– 1.2V
	Transconductance	12 µA/V²
	Linearity of I_{DS} up to V_{DS}	18V
NDMOS	V_t	7V
	BV_{DS}	>100V
	Fully floating	Yes
PDMOS	V_t	– 1.1V
	BV_{DS}	>100V
	Fully floating	Yes

BiCMOS technology with a logic part realized in CMOS, BiCMOS, or bipolar ECL (emitter coupled logic), an analog part using CMOS and/or bipolar, and drivers using bipolar or DMOS (double difused MOS) transistors. The logic CMOS supply voltage typically is compatible with 5V-18V industrial and automotive electronics. For design productivity designers use a cell-based concept complemented with logic synthesis tools for the digital design. The bipolar drivers can reach 30V to 50V, and they output larger currents in a smaller area. DMOS transistors in these technologies can block up to 200V, and they can switch faster. Since Mukkerjee[7] has discussed the broad range of power technologies, we will only consider mixed-signal technologies with 100V to 150V I/O capabilities.

System engineers often use these chips in harsh environments. The chips need a high latch-up resistance and must be able to absorb or block high-energy interference pulses at the terminals. Buried layers and an epitaxial layer help in fulfilling these requirements.

Interface BiCMOS process development. The most recent processes use 3-µm CMOS devices and a 3-µm rule interconnection, which adequately support 100 to 2,500 logic gate integration. The availability of a double-metal interconnection with both metals for signal, or the first metal layer for signal and the second metal for driving purposes, reduces the chip area and improves the design productivity. Decoupling between sensitive analog parts and logic or high-voltage parts is easier with a double-layer metallization interconnection.

We can use up to 40V vertical NPN and lateral PNP bipolar transistors. Also both N- and PDMOS transistors are available; they can block 100V, and the NDMOS has a low resistance when turned on. With these devices we can realize a push-pull output until reaching 100V. Table 3 summarizes the typical DMOS-BiCMOS (DBiMOS) device characteristics. We can now explain the technology and the different layers shown in the cross sections in Figure 1a-c.

A N buried layer implanted in the substrate decreases the collector resistance of the NPN and also the on-resistance of the NDMOS. Together with a P epitaxial layer, it can realize a fully floating CMOS. The buried layer under the N well is also very efficient in improving latch-up resistance. An optimum layer thickness of 15 µm is required to achieve the breakdown voltages of Table 3. The weakly doped N-tub region forms the collector of the NPN and the base of the PNP. The heavily doped N-plug region (Figure 1c) contacts the buried layer and in the CMOS part can increase the latch-up resistance. The N well for the CMOS version is optimized independently from the bipolar layers.

The P-base region serves as a base for the NPN, forming part of the PNP's emitter/collector. It is also used in the NDMOS.

After the base diffusion, the process follows a standard N-well CMOS flow. The NMOS source-drain diffusion serves as emitter for the NPN. The PMOS source-drain contacts the P base.

A typical process needs 14 to 18 masks, depending on the options. The combination of logic, analog signal processing, and driving capabilities permits the integration of a complete intelligent interface system on one chip. The challenge for designers of these technologies is to keep the product cost low enough that costs stay competitive against less integrated multichip solutions.

Several processes covering a broad range of applications are in manufacturing today in Europe.[7,8] The ongoing developments aim at reducing chip size and increasing I/O voltage and current range without adding to process complexity. Innovative device structures or improved layouts based on a better understanding of device physics help designers achieve this goal.

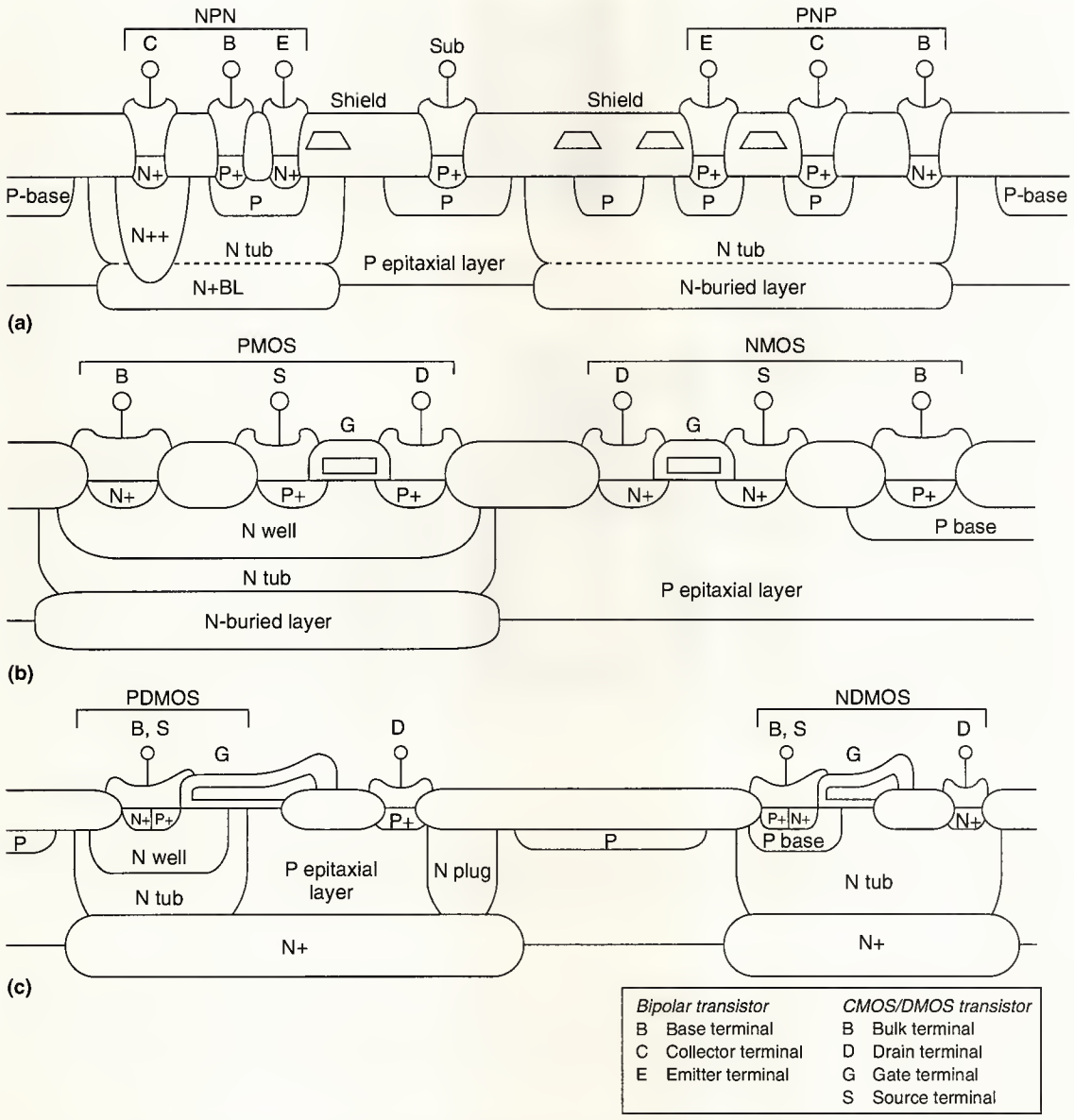


Figure 1. Cross sections of CMOS (a), bipolar (b), and DMOS (c) devices.

Interface BiCMOS cell development. Apart from the specific interface cells, the cell development in an interface BiCMOS process does not differ from the cell development in the mixed A/D CMOS process we described earlier. However, the simultaneous availability in this technology of both bipolar and CMOS transistors opens new, additional possibilities.

First, designers can integrate a very broad range of analog functions simultaneously on one smart interface chip. This range encompasses such functions as precision amplifiers, switched capacitor filters, and A/D or D/A converters for analog signal processing, voltage references and on-chip voltage regulation, power outputs with current and temperature protection, interference protected inputs, and up to 2,500-gate logic for control and/or microprocessor interface. Figure 2 shows a generic example of such an application.

Also, the performance of the functions can be improved

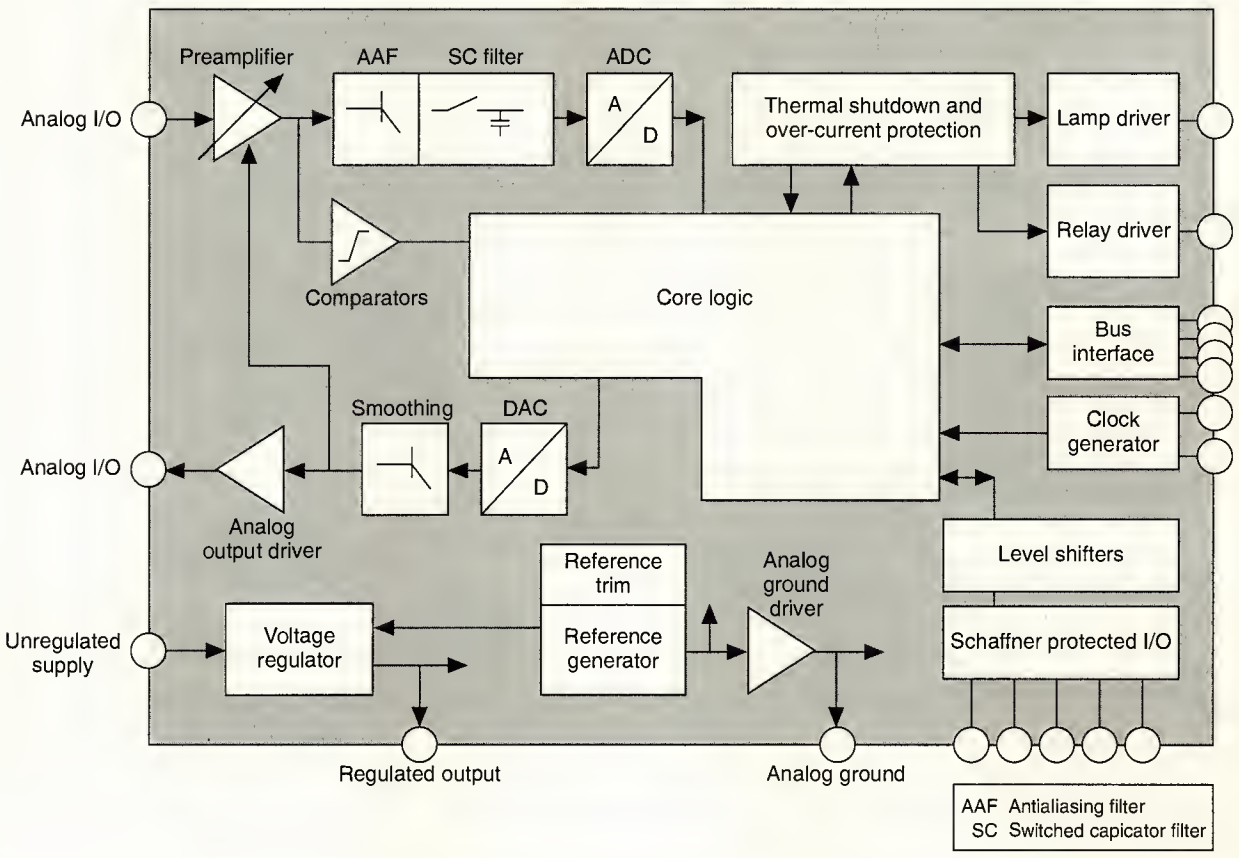


Figure 2. Intelligent interface chip integrating logic, analog signal processing, and output drivers.

by combining both CMOS and bipolar transistors within the same function. Table 4 gives an example of a PMOS input, low-noise, low-input capacitance operational amplifier, and Table 5 describes a low-distortion line driver for mains signaling. A new challenge in this field is the possibility of developing innovative systems by combining the intrinsic high precision of bipolar transistors with the sampled data system approach of MOS transistors.

Finally, the specific interface cells cope with the high-energy interference pulses of the real-world environment. The design of these structures differs greatly from the normal cell design and more closely resembles device engineering. An example is an automotive lamp failure detector circuit. This circuit detects ±10 mV over a sense resistor. The common-mode voltage is the car battery voltage, which can be much higher or lower than the chip supply voltage. The input circuit is also protected against electrostatic discharge and automotive Schaffner pulses, including the +80V load dump.

Table 4. Specification for input capacitance opamp cell (cell library MTC 3100).

Characteristic	Value
MOS inputs	–
Supply voltage	4.5-18V
Supply current	900 µA
V_{os}	< 1 mV
C_{in}	< 10 pF
DC gain	100 dB, typical
Common-mode rejction ratio	70 dB
Power supply rejection ratio	70 dB
Gain bandwidth	2 MHz
Slew rate	1V/µs
Thermal noise	7.5 nV/√Hz
Noise corner frequency	100 Hz

Table 5. Specification of line driver realized in DBiMOS for mains electricity metering.

Characteristic	Value
Supply range	±10V to ±20 V
Output current	< 250 mA, peak
Open loop gain	> 60 dB
Harmonic distortion	< –70 dB; 250 mA, peak
PSRR at 50 Hz	>50 dB
Standby mode	R_{out} > 100 kohms, C_{out} < 50 pF

Low-voltage A/D BiCMOS technologies

Low-voltage BiCMOS technologies are limited to a maximum supply voltage of 12V for the bipolar part and 5V for the digital part. Though this range contrasts with more than the 20V capability just described, it is still a vast domain of technologies that we cannot cover in this overview. Klose recently published a review that included device options for both digital and analog applications.[9]

Pure bipolar processes are still often used for analog designs. The bipolar transistor benefits from a high transconductance, low 1/f noise, a small offset, and a higher drive current on a smaller area. On the other hand the MOS transistor carries a high input impedance and can act as a switch for analog signals. These MOS features formed the base for sampled data analog signal processing as explained for the A/D CMOS process. If we add the availability of large logic cell libraries including dense memories, it then becomes clear that an analog-digital BiCMOS process compatible with a logic CMOS process offers solutions with high performances.

The challenge for this solution lies in the cost-performance trade-off. Compared to the logic CMOS process, four to six extra masks and an epitaxial layer are required. Today A/D BiCMOS processes with minimum dimensions from 1.2 µm to 2.0 µm are in production, and processes with 0.8-µm photolithography are in final stage of development in Europe.

Table 6 lists the typical targets for the 0.8-µm process. This category of relatively expensive processes provides high-performance, high-speed analog capabilities, a maximum analog supply voltage of 12V, high drive currents on chip and at output, and a dense submicron CMOS logic.

Architecture definition is ongoing for a 0.6-µm process. A new dimension here is a high analog performance at a 3.0V supply. We expect to see bipolar transistors, which have already proven analog performance at voltages as low as 1V, become more important here.

EUROPEAN SEMICONDUCTOR COMPANIES and research institutes support a large development effort to produce technologies, cell libraries, and design synthesis tools for analog-digital processes. The developments cover submicron CMOS low-voltage, BiCMOS, and interface BiCMOS A/D processes.

Table 6. Target parameters for typical 0.8-µm analog/digital BiCMOS.

Device	Characteristic	Value
CMOS		As standard 0.8-µm logic process
Vertical NPN	h_{FE}	100
	f_T	> 10 GHz
	BV_{CEO}	> 14V; 12V operation
	V_{Early}	> 40V-50V
	V_{be} matching	< 250 µV
PNP	–	Minimum lateral device
	–	Vertical device for high-speed analog applications requires two more masks
Resistors	–	Precision voltage-independent resistor*
Capacitors	–	Voltage-independent capacitor with good matching performance**

*See hipo resistor in Table 1. **See capacitor in Table 1.

Cooperative European programs in the frames of ESPRIT and JESSI make this effort more efficient. They also forge partnerships between the European semiconductor companies, large system companies, and creative, small and medium-size enterprises to bring timely new products to the market.

References

1. M. Pelgrom, A. Duinmaijer, and A. Welbers, "Matching Properties of MOS Transistors," *IEEE J. Solid-State Circuits*, Vol. 24, No. 5, p. 1433-1440.
2. F. Op 't Eynde and W. Sansen, "A CMOS Wideband Amplifier with 800-MHz Gain Bandwith," *Proc. Custom Integrated Circuit Conf.*, Apr. 1991, IEEE.
3. W. Sansen et al., "Design Tools for Analog Design," *Proc. European Design Automation Conf.*, EDAC Secretary, CEP Consultance Ltd., 26-28 Alongstreet, Edingburgh, EH1 3QH UK, Feb. 25-28, 1991.
4. J. Danneels, "Telecommunications Specific Needs on Technology and Design Methodology," *Proc. European Solid-State Circuits*

Conf., IEEE, Sept. 1991, pp. 7-10.
5. E.A. Geiger, A. Rothermel, and U. Westerkamp, "HDTV: Its Impact on VLSI Architecture and Semiconductor Technology," *Proc. ESSCIRC*, 1991, pp. 47-58.
6. W.P.G. Crooijmans, L. Karman, and R.P. Kramer, "The Realisation of VLSI for HDTV," *Proc. ESSCIRC*, 1991, pp. 59-69.
7. Satyen Mukkerjee, "Advances in Power and High-Voltage Integrated Circuits," *Proc. ESSCIRC*, 1991, pp. 29-36.
8. A. Andreini, C. Conteriero, and P. Galbiati, "A New Integrated Silicon Gate Technology Combining Bipolar Linear, CMOS Logic and DMOS Power Paths," *IEEE Trans. Electron Devices*, Vol. 33, No. 12, Dec. 1986, p. 2025-2030.
9. H. Klose, "BiCMOS—The Technology for Integrating Systems onto One Silicon IC," *Proc. ESSDERC, Microelectronic Engineering*, Vol. 15, Elseviers Science Publishers B.V., Amsterdam, 1991, pp. 501-512.

Edgard Laes is technology development manager at Mietec, where he guides new technology strategy. He has also served as process engineering manager and head of pilot fabrication at Mietec. As a research assistant at ESAT Leuven, he worked on CCD technology development and characterization, lithography (optical and e-beam), and etching. His technical interests include submicron CMOS technology, A/D and intelligent interface CMOS and BiCMOS technologies.

Laes received the master's degree in electronic engineering from the Catholic University of Leuven, Belgium.

Herman J. Casier, currently CAD and design R&D manager at Mietec Alcatel N.V. in Burssels, coordinated the development of the MADE design system and the standard cell libraries for mixed-mode A/D circuits in CMOS and analog BiCMOS technologies. He also helps define new research programs and technologies. In 1980 he cofounded the INCIR, N.V. design house in Leuven, carrying responsibility for custom design.

Casier received an engineering degree in electronics from the Catholic University of Leuven.

Eric Schutz is the business development director in charge of research and development programs at Mietec in Oudenaarde, Belgium, and represents Mietec Alcatel in the JESSI Subprogramme Management Board for Technology. Previously he worked as Mietec's marketing and sales director, and as a member of the Philips design team of ECL 100K for fast bipolar ICs for computer applications.

Schutz graduated as an electromechanical engineer with specialization in electronics from the Catholic University of Leuven.

Direct questions concerning this article to E. Schutz, Mietec Alcatel Headquarters, Westerring 15, B-9700 Oudenaarde, Belgium.

Reader Interest Survey

Indicate your interest in this department by circling the appropriate number on the Reader Service Card.

Low 159 Medium 160 High 161

European Trends in Library Development

The time-to-market and cost constraints of designing VLSI circuits, together with increasing complexity, necessitate a structured design methodology. Such a methodology should be based on an extensive use of libraries of generic components and previously designed macro blocks. Today's libraries, which include objects that were formerly complete LSI implementations, capitalize on generations of designs. However, this reusability depends on a serious effort of design methodology and standardization of CAD environments, macrofunction interfaces, and technology portability.

Jean Pierre Moreau

Joseph Borel

Davoud Samani

SGS-Thomson, Grenoble

Today's integrated circuits are as complex as full systems were several years ago. System complexity has increased too, to take advantage of the more advanced ICs and create more powerful systems. Therefore, design cost and time-to-market restraints still constrain IC design. These constraints can only be met through a structured design methodology based on an extensive use of libraries of generic components and formerly designed macro blocks.

The library is the boundary between silicon design and system design. As time passes, this dividing line rises in terms of the complexity of objects hidden below. Libraries gradually include objects that formerly were complete VLSI chips. This is obviously true for generic macros such as RAMs and ROMs, but it is becoming true also for system components like microprocessor cores or peripherals.

Libraries capitalize on knowledge gained over generations of designs. To do this, successful libraries demand a serious effort at standardization, for at least two basic and very different reasons. The first concerns process evolution. A component's migration from chip status to a library object usually corresponds to a change in technology generation. If the designers of the first version did not plan ahead, the cost of technology migration may be so high that a completely new design may be cheaper. From time to time, redesign may be necessary anyway, because source and target technologies are too far from each other (for example, moving from two metal levels to three metal levels, or even worse from NMOS to CMOS).

The second reason relates to the CAD environment. A library becomes practical when a CAD environment, and now a design automation environment, supports it. CAD is not a more stable world than silicon process. Moreover, different CAD generations from a given vendor are usually not compatible with each other, nor are CAD systems from different vendors compatible, even within the same generation.

Some recent initiatives try to find solutions for the future. Among these is the CAD Framework Initiative, a US-based nonprofit organization gathering several CAD vendors and users to develop architectural and tool communication standards for the new CAD platforms. Another is the European CAD Integration Project, a collaboration project sponsored by the Commission of European Communities to promote CAD standards. The real goal is to make libraries reasonably portable across silicon processes and CAD environments.

0272-1732/92/0800-0043$03.00 © 1992 IEEE

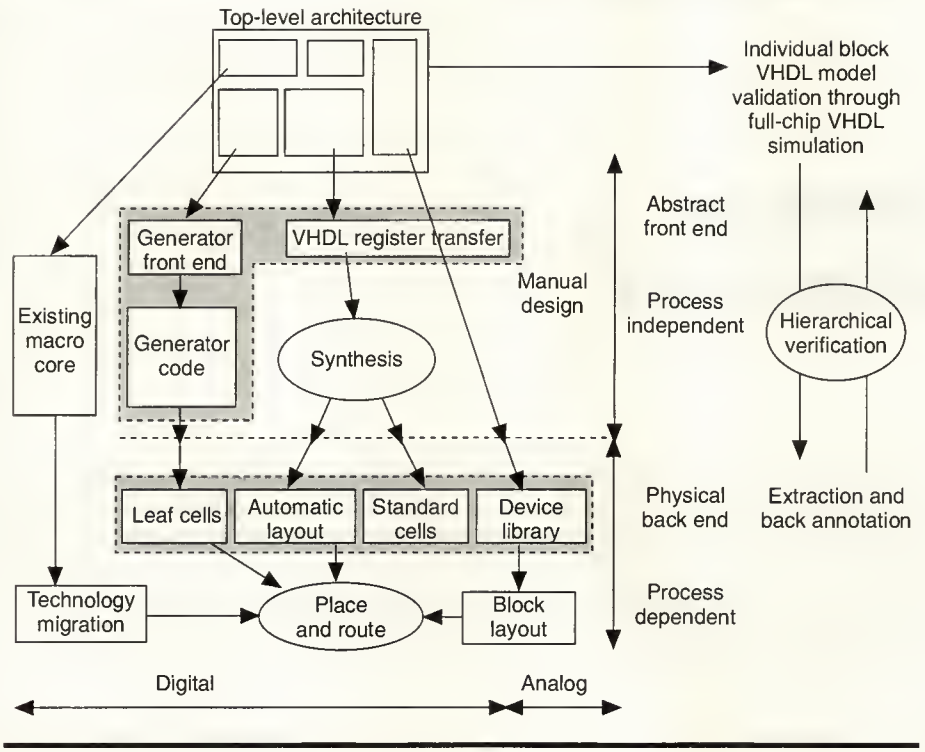


Figure 1. Top-down design strategy.

Libraries for modern design strategies

Ideally, a really useful library relieves the application designer from concerns about silicon device and technology problems. The target is more or less reachable depending on the nature of the application. Fully digital applications are much more likely to achieve design automation than analog or mixed design applications. The level of performance (related to the ultimate capability of the silicon technology in use) is also important. The wider the margin, the easier it is to automate the silicon design process.

Modern design strategies are based on a top-down structured approach. Top-down design starts from formal behavioral specifications (usually validated by simulation) and progressively refines the design by successive partitions of the system into existing chips and chips designed for the new application. Chips are then further partitioned into functional blocks, and functional blocks into simpler ones, down to a level where all leaf blocks are already available (standard cell, for example). If existing chips or leaf blocks are not available in the appropriate technology, they will have to be migrated or designed.

Figure 1 shows a top-down design strategy. The different paths in the decomposition tree may be of different depths. Some will stop very early at a big macro, such as a CPU core, some will descend to simple gate level (standard cells), and others may go down to transistor level.

The latter case necessitates the design of a full custom block. This block must be brought in as a new cell to be validated and then used as a standard cell or block. The presence of such blocks differentiates a full custom design from an ASIC (application-specific IC) design. Everything else can be shared. Even standard cells are becoming more and more popular within the full custom design community, because they save a significant design effort at a very reasonable area cost.

A basic point of the library approach (or, as Hugo De Man, head of the VLSI Systems Design Methodology Department at the Catholic University of Leuven, calls it, the "meet-in-the-middle strategy,") is that everything below the line is, at least for an important part of the design process, hidden to the application designer. The application designers must rely on the abstract views of the library objects (such as behavioral models or timing information). The library designer must provide this information and properly validate and characterize to guarantee the parameters used at system level.

We call library views related to the application engineer's work front-end views. Views related to the physical implementation are referred to as back-end views. Timing data excepted, front-end views are technology independent. Their portability problems relate only to CAD tools. Each tool needs its own format and therefore a dedicated view. The back-end views are not only CAD dependent, they are also process dependent.

Portability of library views

A portable library can be installed at reasonable cost across different CAD platforms. Front-end views require the highest flexibility (because of the variety of users), while back-end views may exist on a smaller number of systems. However, the problem is basically the same in both cases and relates essentially to the existence of CAD tool independent standards for describing the library information.

Fortunately for application engineers, the situation is far better for front-end views, where a significant standardization effort is beginning to provide results (see next section). The situation is not so advanced for back-end views, which

include the layout view. Two reasons explain this. First, less people are involved (the silicon vendors), and those people may have found some interest to define and keep for themselves their own standards. Second, the problem is intrinsically more difficult, because the way back-end views are described is the key for migratability. This may also be a reason why work in this field has been confined to silicon design houses for years.

We explain various methods for migrating layout from one technology generation to the next later in this article. They range from simple optical sizing to full redesign, depending on the differences between source and target technologies and performance constraints, which may become considerable when full redesign is necessary.

Front-end to back-end interface. The major steps for ASIC design are

- definition of a behavioral model,
- definition of a structural model,
- functional simulation,
- partitioning into simpler functional modules,
- early floor planning (mainly for full custom design), and
- prelayout simulation.

In the traditional ASIC approach, the interface is a netlist of library elements. Placement and routing take place in the back end under the responsibility of the ASIC vendor. However, for very complex circuits requiring more than standard cells, there is a tendency to drive the logical partitioning with layout, or at least topological information, using a family of tools known as Design Planner.[1] In this case, we want some visibility of the back-end world. Usually we get this through abstract views (sometimes called place-and-route views), which give information on the size and terminal position of the library elements. The problem is that each place-and-route tool requests a specific type of data according to its own capabilities and limitations. This is much more than a simple problem of format; each tool requires its own semantics.

Toward digital standards. Designers widely recognize the need for standards, which are now appearing in the front-end world. The development of real standards depends on two conditions: The semantics of the problem should be clear, and the user community should be large enough to guarantee an unquestionable return on investment. This is clearly the case for most of the front-end representations and especially for functional representations, made up of a mix of behavior and structure.

Traditionally, designers have represented their design through a description suitable for simulation. However, there are as many languages as there are simulators. This could be for good reasons (such as, the electrical level can hardly be described with the same conventions as higher levels) or for bad reasons (each simulator vendor has developed its own language and considers it superior to all the others).

VHDL. The VHDL initiative[2] approached the problem in the reverse way, saying: Let's first define a hardware description language with all the features and internal consistency of a high-level programming language and make it a standard. Then simulator vendors will be free to develop their own engine for that language. Ideally, every construct of the language will be unambiguously defined (as far as behavior is concerned), and by definition all VHDL simulators will give the same results for the same circuit and description stimuli. VHDL can be a common reference for all people working on a design and, in particular, between system designers and ASIC vendors.

Although this uniformity is a real possibility, the situation is not this simple, for a few reasons. First, some ambiguities remain in the standard at present. Second, most VHDL simulators do not support all the features of the language. Third, VHDL lacks some features that would make designers' lives easier. For example, VHDL cannot automatically back annotate simulation models with data such as parasitic capacitances extracted from real circuit layout.

Nevertheless, VHDL's availability represents considerable progress compared to the past. As it is, VHDL allows designers to write simulator-independent specifications and models and to structure libraries by keeping only one copy of the data common to several variants (for example, according to technology variants or generations). See Figure 2.

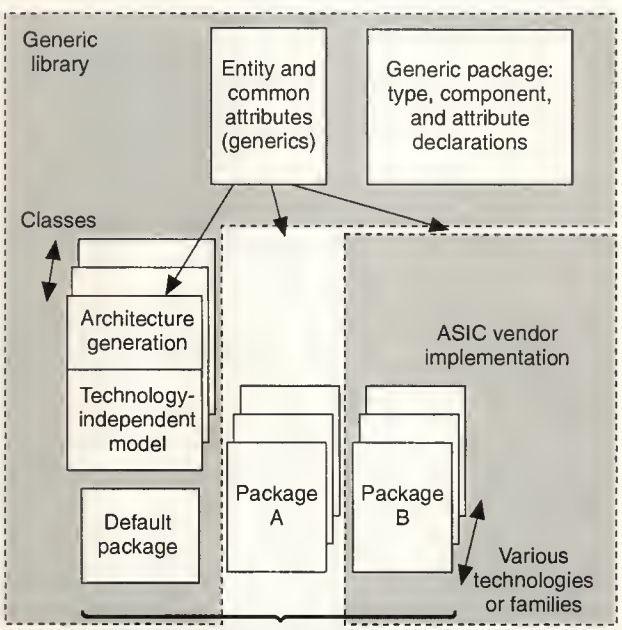


Figure 2. Libraries structured by variant.

More generally, VHDL codes front-end information (including power, timing information, and area) by attribute. This means a designer can build a reference library using only VHDL and a VHDL database. Then, through procedural interfaces, tool-specific representations can be automatically derived, simplifying library maintenance and tool exchange. See Figure 3.

Steven H. Kelem proposed this approach at a VHDL user's workshop in October 1990. One drawback is that the standard defines only the syntax of VHDL attributes. Users must define the meaning of particular attributes and provide tools to interpret them. To be really useful, this approach should generate another layer of standardization.

EDIF. The Engineering Data Interchange Format represents the other major achievement in digital standards. EDIF is a neutral format, originally promoted by a group of CAD vendors to make circuit netlist descriptions exchangeable between different CAD systems. EDIF standard groups are very active, particularly in Europe.

EDIF is sometimes presented as a rival of VHDL, which is not the case. EDIF is first an interchange format and then a description tool for those representations for which a semantic layer has been agreed to on top of the format. This is essentially the case for netlists. A design made of standard cells can easily be ported from one CAD system to another through an EDIF representation of the netlist. However, this is seldom as straightforward as it could be, because only the format is really part of the standard; how it is used may be subject to interpretation. It is usually not too difficult to convert an EDIF netlist for Cadence into one suitable for Mentor, but some corrections are needed to get through.

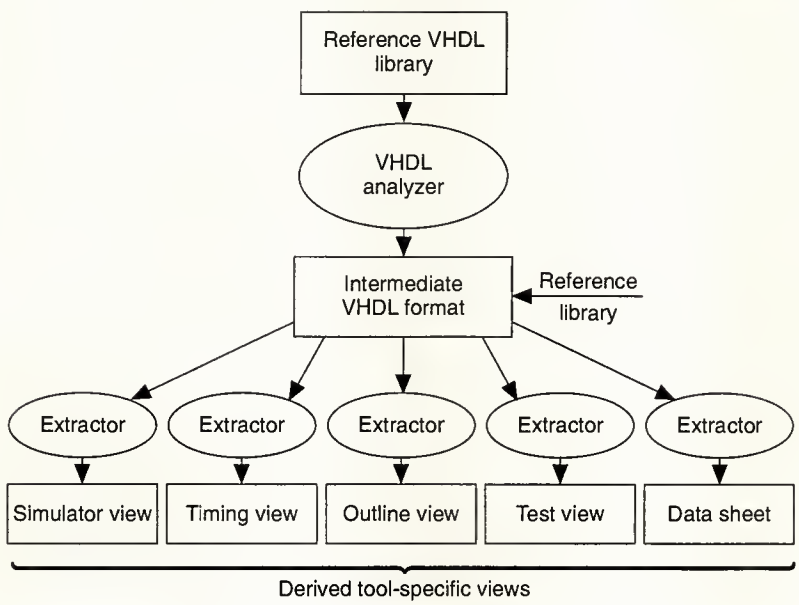


Figure 3. Tool-specific representations derived from a VHDL reference library.

The analog case. Analog libraries share many views with digital ones and therefore can use the same standard formats. However, information related to behavior is more difficult to express, and standards and tools developed for digital are not really suitable for analog.

Researchers are exploring two approaches. The first, favored by those who work with mixed design, is to extend digital tools to analog. This extension would introduce new features for describing analog behavior into VHDL and develop analog functional models that could run in a mixed environment.[3,4]

The other approach declares VHDL to be too crude to deal with analog requirements, whatever a "VHDL extension" may be. This is clearly the case for designers dealing with high-performance, pure-analog design. They don't use a predefined library, with the possible exception of a library of elementary devices or structures, such as current mirror or voltage references. We give some complementary details later.

A combination of the two approaches may be the right path. The first will never be accurate enough to optimize the pure-analog part of a design. However, the VHDL extension is necessary to support a global simulation of a circuit including both complex digital processing and analog processing.

Toward layout automation

The cost of physical layout of library elements continues to rise because they continue to grow more complex. This makes migratability (the capability of a library to be moved easily to a new process) an important feature.

The problem is not one of representation format; rather it is a problem of layout strategy. It makes sense to define an associated standard format only if good strategies are used.

Note also that a migration that keeps the original transistor-level schematics needs only to be recharacterized in the target technology. If the schematic is somewhat modified, a full transistor-level validation becomes necessary. Automatic tools are now available to support this operation.[5]

The ultimate goal is an automatic layout generation. Most researchers base their approach to this problem on a hierarchical partitioning of complex layouts into smaller pieces down to physical leaf cells.

Choosing a strategy for automatic leaf cell assembly. Designers use three main approaches to assemble leaf cells:

- procedural code (tiler), for regular arrays;
- application-directed placement plus automatic routing and optional compaction (usable for semiregular arrays such as data paths); and
- automatic placement and routing plus optional compaction, preferred for quasi-random structures.

Of course, the structural partitioning may take more than one step. Each partitioning subtree may use a different strategy. In this case, we would use the second or third strategy to assemble the subtrees. In any case the assembly tool must be able to handle leaf cells in an object-oriented way; that is, the assembly code should be independent of the size of the leaf cells.[6]

The ultimate goal is an automatic layout generation.

The second and third approaches can make use of standard tools, but this is not always the case, particularly for routing. Most available routers are designed for ASIC standard cells. Leaf cells in a dedicated block are usually designed according to some specific rules to make this particular block more compact. For example, designers use abutment as often as possible, or they use multiterminal cells to bring extra degrees of freedom for routing in a data path structure. Dedicated routers may be used, and rather than one standard router, a full routing tool box is needed.

The first approach uses an assembly code that may be anything from C or a CAD system extension language (such as Cadence's Skill) to a dedicated layout language (such as Mentor's L or Styx, SGS-Thomson's proprietary language built as a Lisp extension). Designers normally use this approach only for regular arrays, so there is almost no routing process. Connections are made by abutment or, less frequently, use local routers, such as a river or a channel. Those routers are usually dedicated to the procedural layout environment for better efficiency.

Leaf cell design. Technology dependency is concentrated at leaf cell level. It is not unthinkable to handcraft a set of leaf cells at each technology generation. However, there are some possibilities for improvement, and in some cases it may be possible to automatize the generation of leaf cells themselves.

Migrating "polygon-pushing," hand-crafted leaf cells. Polygon pushing, though regarded by some as practically prehistoric, is still commonly used. This migration technique consists of processing a layout originally drawn for technology A (say, 1.2-micron CMOS) to match as closely as possible the rules for a process B (say, 0.8-micron CMOS). The migration requires two complementary actions, optical shrinking and polygon sizing. New mask layers can be generated through "logical" combination of already existing levels.

This approach has two drawbacks. First, the final rules are seldom the ultimate allowed by the target technology. The situation may be even worse if source layout was already the result of a migration from an older technology. The second drawback is that the designer has no control over the electrical performance of the result. Power and access resistances may be too high. From migration to migration, the overall performance compromise deteriorates and technology capabilities are wasted.

To solve those problems, some researchers have proposed a more sophisticated approach based on circuit recognition. First, an electrical extractor is run on the original layout, resulting in an electrical schematic following the layout's topology. This schematic is also a kind of symbolic representation of the layout that can be taken dimensionless, that is, the really important information relates to relative positions, not size. Designers can then expand this dimensionless layout according to the rules of the target technology and target performances. As a component (and not only as a meaningless set of polygons), each element of the design can be resized to the optimum value. This approach directed the development of the Desire system marketed by Dosis.

Procedural leaf cells. Sophisticated procedural layout languages (such as L, Styx, or Philips/S3's Mod Gen) support a completely rule-independent leaf cell coding. This however is rather theoretical. In practice, a true rule-independent code is impossible to write for the general case. It is only possible to migrate to similar technologies, for which the most constraining design rules follow the same order. If the order is not the same, the code must include complementary test and most often will have to be patched and revalidated for the new technology.

In fact, the approach is efficient only for very simple leaf cells, such as the primitive elements for the symbolic approach explained next. Rather than leaf cells, we should talk of "parameterizable generic components." Those components are then associated in cells using a symbolic and compaction approach.

Symbolic layout approach. Symbolic layout is based on predefined elementary components that may be as simple as a contact via or transistor. Symbolic elements may be fixed (such as a fixed-size via between metal 1 and metal 2), parameterizable (a sizable transistor), or path symbols (a connection wire). Those components are assembled on a drawing grid, which may be fixed or virtual. In the latter case, compaction software implements each component as close to its neighbor as allowed by the process design rules. Effi-

cient systems should include a capability for monitoring compaction, for example, to match a particular position or pitch.

Symbolic layout is highly migratable, because technology dependency is only at the elementary component level. The fixed-grid approach—by far the simplest—doesn't use the ultimate technology resource. Cell comparison gives a difference between a fixed grid and the compacted symbolic of about 15 percent. However, this may not be so significant at circuit level, because a fixed grid enforces an extra regularity that may ease abutment and other area-saving connection techniques.

We emphasize reusing the system design effort and automating the layout part as far as possible.

Automatic generation of layout. Modern three-layer metal CMOS technology makes automatic layout generation—the ultimate goal of libraries—realizable for digital design. This approach concentrates the technology dependency in a very small number of elementary components, say a P-channel and an N-channel transistor, as in symbolic layout. Those transistors may have up to four degrees of freedom (including size, one or two coordinates, and orientation) or no freedom at all. In the latter case, the topological target is like a sea of transistors.

The more degrees of freedom available, the more difficult the problem becomes. In practice, the freedom is limited. Transistors are placed in parallel rows of N and P transistors, similar to a sea-of-gates topology. However, provision for sizing may remain and transistors are grouped in clusters.

An automatic cell layout system, such as our company's Gencell,[7] begins with a transistor netlist, which may be produced by a dedicated electrical synthesis tool. The system produces the layout according to a particular strategy, including component sizing to match speed and power criteria. To migrate to another technology, a designer needs only to run the tool with a suitable library of elementary components. Compared to a pure standard-cell approach, automatic layout with optimum sizing offers an extra level of flexibility that may be of interest for critical designs. Also, the front-end synthesis may target complex gates (complex, inverting gates, implemented in only one layer) that may be smaller than a combination of standard cells.[8]

Overall, automatic layout with component sizing may give a significantly better compromise between power and area than do standard cells. The result can be used as either a hard macro or a specific block used for only one particular application. In the first case, the designer must run a full characterization to include the new block in the library. In the second, the designer may choose to keep the generated block at the transistor level and use a multilevel resistive-switch simulator to check the timing performance of the chip.

Migrating complex libraries across different technology families. Functional abstraction and logic synthesis tools[9] for automatic layout generation open the door to more ambitious migration schemes. For example, migration of a complex NMOS circuit core could be reused as a macro in a CMOS environment.

Researchers have already experimented and proved this technique efficient. The first step in the technique is to generate a register-transfer-level VHDL model suitable for synthesis. This may require partitioning the source circuit.

The starting point is the switch-level simulation model usually developed for test pattern generation. This model mixes switches, gates, and simple macros (ROM, for example). The abstraction tool extracts the register-transfer functionality of the switch networks. It does this partly by identifying canonical forms and partly by recognizing forms (for example, latches).

Next, we can input the resulting model to a synthesis tool able to manage both the glue logic and the ROM-like macros. The target library is a CMOS standard cell library that includes the necessary generators for regular arrays. The synthesis tool generates a netlist that is fed into a place-and-route tool, which generates the final implementation.

This is far from a push-button process, since most of the tools are still experimental and need to be tailored to the application. (This is particularly true for abstraction.) However, it opens new routes to make previous design effort reusable. Note that we emphasize reusing the system design effort and automating the layout part as far as possible.

The analog case. Analog is highly sensitive to process parameters. Designers must size individual components very carefully and resize them for each new technology. Also, component placement must conform to specific constraints to maintain symmetries, keep cross coupling low, and so on.

As a result, analog specialists usually draw analog cells by hand. Since simple global sizing is probably much too crude for process migration, components must be individually resized and the full cell redrawn for each new process, even if the differences are small.

Several researchers have attempted to automate this process, and considerable work continues. The technique is based on a two-step approach and relies on electrical and topological templates.

The root library contains a set of functions defined at the structural level (the electrical template). This set of functions is the knowledge base of a tool capable of sizing each com-

ponent of the template according to a set of functional specifications. Then another tool draws the individual components at the right size, places them, and connects them according to the topological template.[10-12] In such a process, the analog (human) expert must define the initial templates. Then many variants can be generated for a very low cost.

Let us consider some other details. The most important problem in synthesizing analog circuits (unlike digital circuits, in which only timing is of concern) is to satisfy the very wide spectrum of input specifications. Analog specifications are also analog. Casting a glance into the full set of specifications for an operational amplifier, shown in Figure 4, clarifies this specific fact of analog design and its constraints.

We can divide the various approaches to analog synthesis into four separate classes.

- High-level analytical approaches have been developed for switched-capacitance circuits. (Our company, for example, has developed a switched-capacitance filter compiler.) These approaches are based on mapping the desired transfer function to a sequence of circuits with known behavior, such as biquadratic filters.
- The second class, parameter-preprocessing-based methods, relies on parameterizing part of the circuit layout. These methods are useful for obtaining partially parameterized module generators. They carry out a correct manual layout, identify its parameterizable sections, and then modify the appropriate portion of the geometric representation. Despite its real advantage as a standard cell library approach, this process is lengthy, error prone, and difficult to generalize. Moreover, it is not a viable method to obtain technology and design-rule independence.
- A third class relies on developing routines specialized and optimized for a given analog function. In recent years several computer programs of this kind have been developed (for example, Adam shown in Figure 5 on the next page). In each of these programs, the user selects a particular circuit template from the program's library, and the program helps the user determine numeric values for that template's parameters. The program uses design equations associated with that template to find values satisfying the user-specified parameters and behavior constraints, possibly also optimizing an objective function.[13]

  We can generate the mathematical equations describing circuit behavior in two ways: analytically (derived manually by the analog designer) or numerically. The analytical approach is more efficient than the numerical and is, in general, the preferred method. Unfortunately, automatically generating design strategies for new electrical templates is very difficult, and thus the analytical approach will likely remain a fixed-template technique for some time.

DC analysis
- DC currents and voltages for all branches and nodes
- Small signal parameters for all transistors
- Common-mode input voltage range, output voltage range and maximum output current

AC analysis
- AC resistances and capacitances on all nodes
- Gain and bandwidth versus frequency and biasing current
- Slew rate and settling time versus load capacitance
- Input and output impedances versus frequency (for open and closed loop)

Second-order characteristics
- CMRR (common mode rejection ratio), PSRR (power supply rejection ratio) and equivalent input noise (voltage and current) versus frequency
- Offset voltage versus common-mode input voltage
- Drift versus temperature, time, and supply voltage
- input bias current and input offset current

Figure 4. Specifications for an operational amplifier.

On the other hand, the numerical optimization approach, although less efficient, is more flexible. This method does not restrict the user to a fixed, predetermined set of templates. It can accept new templates from the user and generate the appropriate set of design equations for that template. The task remains how to create automatically designed plans from obtained equations.

A group led by Willy Sansen at Catholic University of Leuven says it has developed a system that can perform this task. According to the group, the Ariadne system uses the Donald equation manipulator to design plans from obtained equations (see Figure 6).[14] Donald manipulates the mathematical equations, obtained through the symbolic simulator Isaac, to arrange them to meet input specifications. SPICE simulations in the loop verify the accuracy of the results after optimization. If specifications are not fully met, the user must modify the initial choices.

- The fourth class, the knowledge-based approach to analog synthesis (such as S3's Midas), combines fundamental rules for various devices with network laws to describe a device's behavior. Given a goal, designers use the rules to dynamically decide which components to use and how to interconnect them. This method cannot take ad-

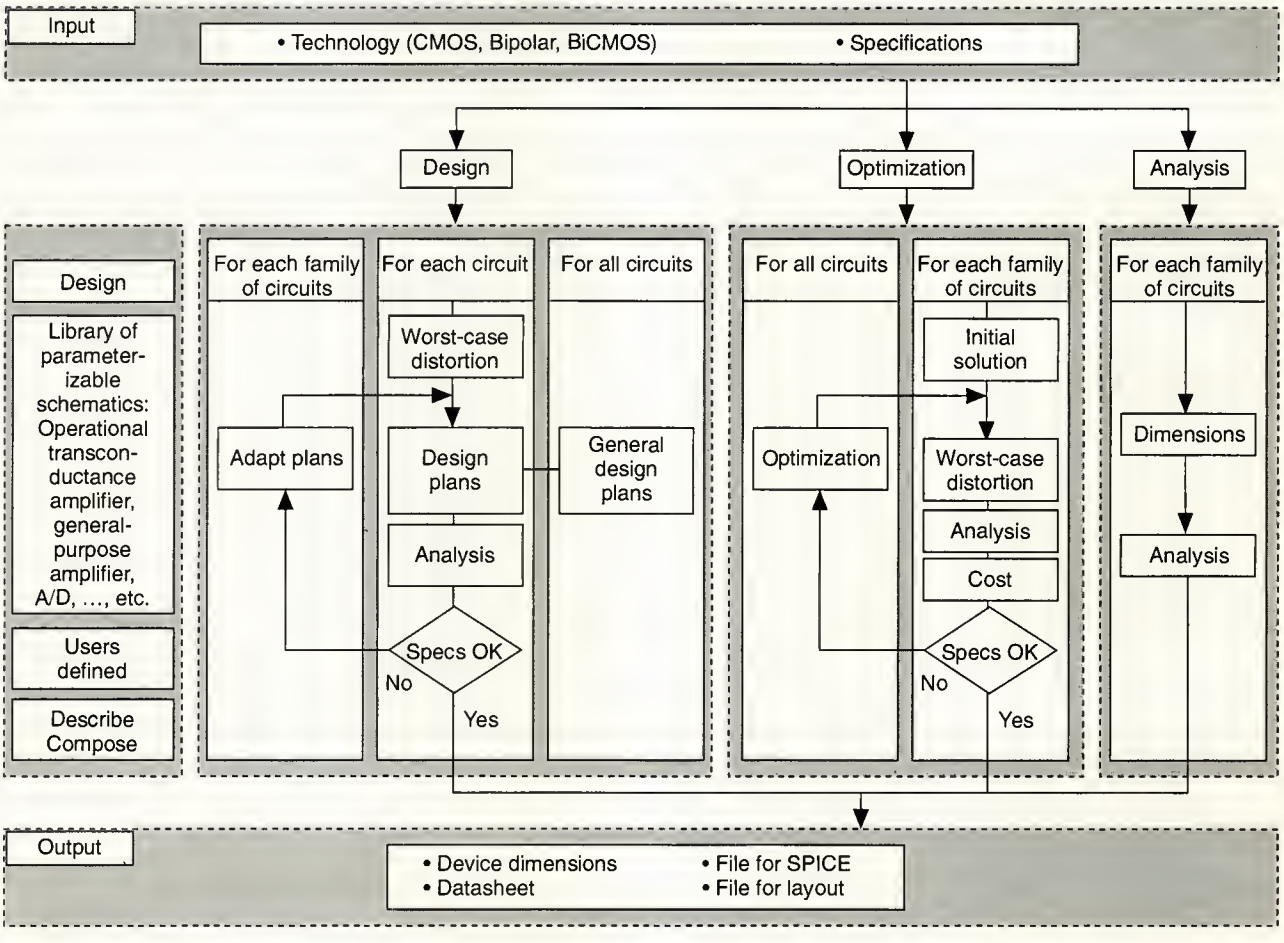


Figure 5. Overview of the Adam system. (Courtesy of Centre Suisse d'Electronique et de Microtechnique.)

vantage of specific templates and their known usefulness. Since this approach manipulates primitive elements and an associated algebra dynamically, it is difficult to use for designing complex circuits.[15,16]

To ensure a high yield, analog silicon compilers should, during their design cycle, take into account temperature, power supply, and technology parameter fluctuations. We can do this by using a circuit analyzer with worst-case capabilities in an iterating design loop. However, this method is less than ideal since it takes considerable CPU time to achieve the result, which may in fact be unrealistic.

Another solution is to predistort the specifications for the typical case in such a way that the circuit meets specifications beyond worst-case conditions. (Adam uses this method.) This process runs on a flat schematic for equation handling, which is a prerequisite for predistortion of input specifications.

Once these equations are obtained, a new set of specifications at reference temperature is computed using the minimum supply voltage and "slow" electrical parameters to size the circuitry. We cannot apply this method to hierarchical approaches, such as knowledge base approaches, because the circuit should be known in its flat form before worst-case handling of equations.

Mixed design. For an even better integration standard, we can mix analog and digital functions on the same chip. This approch necessitates careful design practices and adequate technologies. From the technology standpoint, designers usually consider two routes: multipurpose CMOS and BiCMOS. Multipurpose CMOS targets a slightly different performance compromise than pure-digital CMOS, because higher voltages must be handled for analog cells and, possibly, for EPROM and EEPROM cells. The two compromises (low voltage for pure digital, higher voltage for analog) also exist for BiCMOS.

Experience shows, however, that in digital applications BiCMOS provides high drive buffers for heavily loaded internal clocks and pad drivers. The digital section is always better when made with pure CMOS circuits. So, CMOS digital libraries, originally made for digital CMOS technology, should be migratable to multipurpose CMOS or BiCMOS.

However, this migratability constrains the process definition and rules for making digital libraries reusable (perhaps with no more change than a recharacterization) in a multipurpose CMOS or BiCMOS process from the same generation. Direct reuse may not always be possible and some migration action may be necessary. In any case, process engineers do their best to keep the migration path within the limits of a simple resizing. Our company systematically uses this process development strategy to make all CMOS digital macro blocks easily reusable in BiCMOS.

System-level libraries: functional standards

Besides portability and migratability, we should introduce the concept of interoperability. It is the natural consequence of the rising complexity of the library blocks, which include microprocessor cores, microprocessor peripherals, and dedicated communication interfaces. Library designers must now concern themselves with interface specifications, communication protocol definition, and bus specifications.

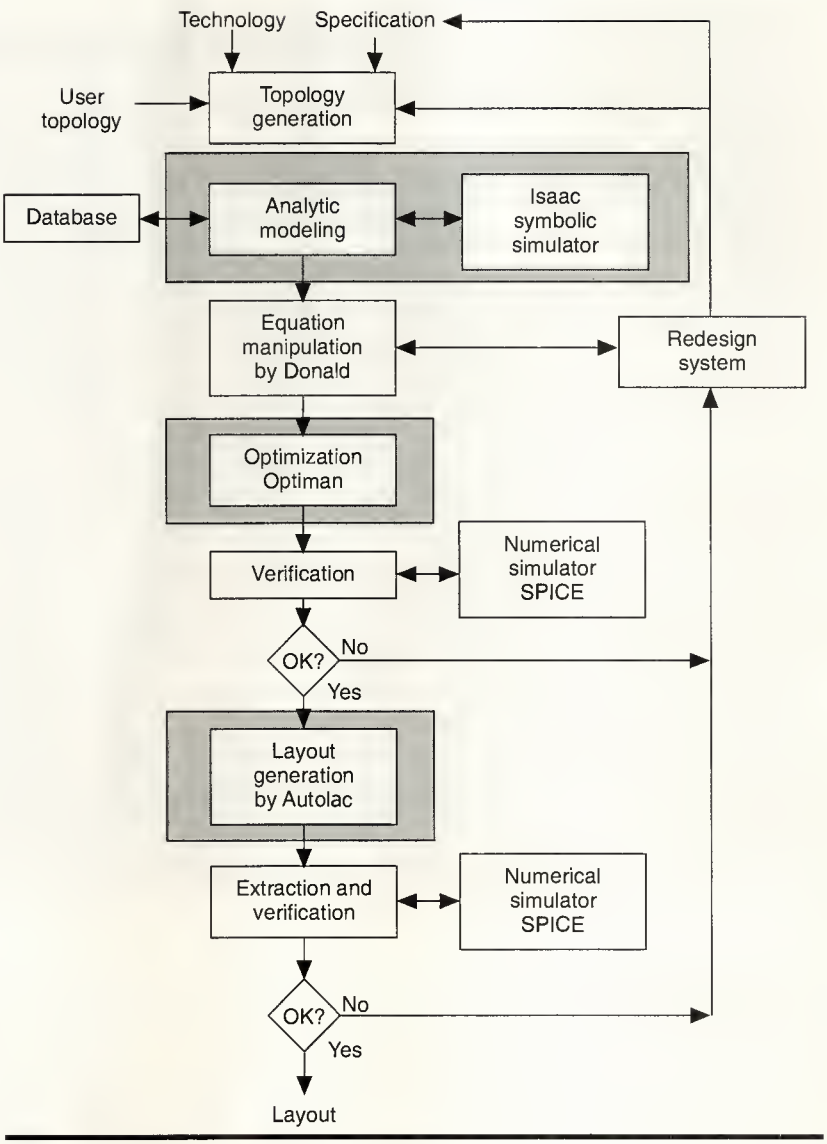


Figure 6. Design flow with Ariadne.

THE DESIGN COMMUNITY has invested considerably in the library field. Most researchers admit that such an investment cannot be repeated for each process generation. The ever-increasing complexity of library elements and the constraints of bringing new technology to market in the shortest possible time demand that this investment be reusable to a large extent from one technology generation to the next.

However, for lack of standards and methodology, most of these investments are today hardly reusable in contexts that differ from their creation. European research recognizes this problem as a major one and aims at building more structured approaches based on internationally recognized standards and modern design methodologies. (See box on the next page.)

New tools are being developed to automatically characterize new designs. Researchers are experimenting with tools for layout conversion or automatic generation.

When automatic layout generation is achieved, we will see a significant convergence between full-custom and semicustom ASIC design. A unified, structured design methodology will emerge, giving designers and system architects more flexibility and freedom in their activities.

European library research programs

Two major library research programs are under way in Europe.

ESPRIT. In the European Special Program for Research in Information Technology, library research is driven by the goal of supporting ASIC design methodology or by process research.

The Integrated Design and Production System (IDPS) aims to establish the foundation for promoting standard design methodology and basic reference libraries. We expect very few application-specific cells from IDPS. Rather, IDPS will provide the first level (standard cells and general-purpose generators such as ROMs, RAMs, multipliers, and the data path) and the overall methodology, based on a top-down VHDL approach.

IDPS results will be embedded within semicustomized products offered by the vendor members of the project and available to system designers and research institutions, including those in the Eurochip association.

More specific libraries are coming from the Advanced Programmable Building Blocks and Candi (BiCMOS development) projects. Both projects commit to provide interfaces compatible with IDPS rules and standards.

JESSI. Research in the Joint European Submicron Silicon program is driven by the so-called Euro projects. These projects are application oriented (such as HDTV or high-security systems for cars), and they usually cluster a system development project with an application-oriented CAD project (such as Analog Expert or synthesis).

The JESSI Submanagement Board for Applications is discussing development of application-specific libraries. The rationale is to strengthen each application cluster with a special library project, to make results reusable in different contexts but in the same field of application.

The Analog Expert Design System project is also of major importance for the future of analog library design. It seeks a flexible environment for analog libraries and mixed-systems design. The main areas addressed are

- a high-level description language suitable for analog,
- synthesis of a hierarchical netlist for analog building blocks at transistor level from high-level description using expert knowledge,
- conversion of transistor netlist into layout,
- mixed digital/analog, multimode (functional to transistor) simulation,
- back annotation of parasitic layout elements in simulation models, and
- a library of analog circuit templates and synthesis rules.

Acknowledgments

The Commission of European Communities through the ESPRIT programs and the JESSI initiative sponsored most of the works used to prepare this article.

References

1. C.H. Masson et al., "Object-Oriented Lisp Implementation of the CHEOPS VLSI Floorplanning and Routing System," *Proc. 28th ACM/IEEE Design Automation Conf.*, IEEE Computer Society Press, Los Alamitos, Calif., 1991, pp. 259-264.
2. R. Lipsett, C. Schaefer, and C. Ussery, *VHDL: Hardware Description and Design*, Kluwer Academic Publishers, Boston, 1988.
3. B.R. Stanisic and M.W. Brown, "VHDL Modeling for Analog Digital Hardware Designs," *IEEE Int'l Conf. CAD Digest of Technical Papers*, CS Press, 1989, pp. 184-187.
4. C.M. Kurker et al., "Development of an Analog Hardware Description Language," *Proc. IEEE Custom Integrated Circuit Conf.*, CS Press, 1990.
5. K. Anshumali, "ACC: Automatic Cell Characterization," *Proc. Euro-ASIC Conf.*, CS Press, 1991, pp. 204-209.
6. A. Compan et al., "Genview: A Portable Source-Level Debugger for Macro-Cell Generators," *Proc. European Conf. on Design Automation (EDAC)*, CS Press, 1991, pp. 408-412.
7. R. Jamier and J. Frehel, "Automatic Layout of CMOS Cells," *Logic and Architectural Synthesis*, Int'l Federation for Information Processing, North-Holland, Amsterdam, 1991, pp. 265-274.
8. M. Crastes et al., "A Technology Mapping Method Based on Perfect and Semi-Perfect Matching," *Proc. 28th DAC*, CS Press, 1991, p 93-98.
9. Ton Kalker, "Formal Methods for Silicon Compilation," *Proc. EDAC*, CS Press, 1991, pp. 395-400.
10. J.D. Conway and G.G. Schrooten, "An Automatic Layout Generator for Analog Circuits," *Proc. EDAC*, CS Press, 1992, pp. 513-519.
11. J.M. Cohn et al., "KOAN/Anagram II: New Tools for Device-Level Analog Placement and Routing," *IEEE J. Solid State Circuits*, Vol. 26, No. 3, 1991, p 330-342.
12. S.W. Mehranfar, "A Technology-Independent Approach to Custom Analog Cell Generation," *IEEE J. Solid State Circuits*, Vol. 26, No.

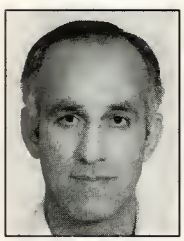
3, 1991, p. 386-393.
13. M.G.R. DeGrauwe et al., "IDAC, an Interactive Design Tool for Analog CMOS Circuits," *IEEE J. Solid State Circuits*, Vol. 22, No. 6, Dec. 1987, pp. 1106-1117.
14. K. Swings and W. Sansen, "Donald: A Workbench for Interactive Design Space Exploration and Sizing of Analog Circuits," *Proc. EDAC*, CS Press, 1991, pp. 475-479.
15. L.R. Carley and R.A. Rutenbar, "How to Automate Analog IC Designs," *IEEE Spectrum*, Vol. 25, No. 8, Aug. 1988, pp. 26-30.
16. B.J. Sheu, A.H. Fung and Y.-N. Lai, "Knowledge-Based Approach to Analog IC Design," *IEEE Trans. Circuits and Systems*, Vol. 35, No. 2, Feb. 1988, p. 256-258.

Jean Pierre Moreau directs European Collaboration Programs at SGS-Thomson Central Research and Development, Design Automation and Integrated Systems in Grenoble, France. His technical interests include libraries and CAD tools for design automation. He previously led the company's semiconductor CAD development team.

Moreau received his Engineer's Degree in electrical engineering from Ecole Superieure d'Electricite in Paris and his PhD in semiconductor physics from the University of Dijon. He is a member of the Societe des Electriciens et Electroniciens and IFIP Working Group 10.5.

Joseph Borel is vice president of Central Research and Development in SGS-Thomson's Design Automation and Integrated Systems organization. He was previously technical director for Thomson Semiconductors.

Borel received his Engineer's Degree in electrical engineering from Ecole Superieure d'Electronique et de Radioelectricite in Grenoble and a PhD in physics from Grenoble University. He is a senior member of IEEE and served as associate editor of *IEEE Transactions on Solid State Electronics*. He is also an associate member of the application committee of the French Academy of Sciences and a member of the JESSI Submanagement Board for Applications.



Davoud Samani manages the Analog and Mixed Design Solutions team, also in SGS-Thomson's Design Automation and Integrated Systems organization. He has also worked in analog and mixed circuit designs.

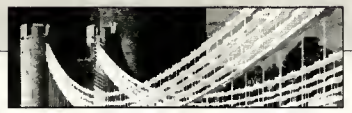
Samani received his BS degree from Pahlavi University in Chiraz, Iran. He received his Engineer's Degree and PhD in microelectronics from Ecole Superieure d'Electronique et de Radioelectricite in Grenoble. He is a member of the steering committee of the Network for European Analogue Research.

Direct questions on this article to Jean Pierre Moreau at SGS-Thomson, BP 217, F-38019 Grenoble cedex, France; or via e-mail at jpierre@dt8600.st.it.

Reader Interest Survey

Indicate your interest in this article by circling the appropriate number on the Reader Service Card.

Low 156 Medium 157 High 158



European Activities for EDA Standardization

Current European activities for standardization in the field of electronic design automation concentrate mainly in the areas of system description languages (VHDL), data interchange formats (EDIF), and development environments (frameworks). As basic work for the support of these three areas, information modeling (Express) and standard interoperability also became important areas of investigation. These activities started in 1986 under the ESPRIT program. In 1991 a JESSI Subprogramme Application launched the European CAD Standardization Initiative (ECSI), using the ESPRIT activity as the right basis, and giving this research activity the needed industrial umbrella.

Anton Sauer

Siemens Nixdorf

Jean Pierre Tual

Bull S.A.

Robin La Fontaine

Monsell EDM

The need for standardization in the field of electronic design was well recognized from the very first use of computer-aided design tools. The problem arose as electronic designers gradually became more and more connected to some domains outside their main field of expertise.

At first, designers were mainly linked to factories responsible for manufacturing integrated circuits and printed circuit boards or systems and to organizations immediately surrounding their company, such as the drafting shop for PCB plots, and the design automation and component library departments. In the current phase, electronic designers also need to establish smooth links with other areas of automation such as 3D CAD for mechanics and computer-aided software engineering for software development.

The search for ways to integrate different tools and different organizations via efficient interfaces initially drove the need for standards. Later, new areas developed such as computer-aided concurrent engineering (CACE),[1] which now acts as a major driving force for standards. The need for standards in the electronic design field is most crucial and will influence the development of applications in all other segments of concurrent engineering. In the electronics field standards need to be established in three main areas:

- *data interchange formats* (for example, EDIF, the Electronic Design Interchange Format[2]);
- *system specification and description languages* (for example, VHDL,[3] the VHSIC Hardware Description Language[4]); and
- *integration platforms* (frameworks).

The ECIP

Six years ago major European companies, together with important institutions and universities, started activities to develop electronic design automation standards. They wanted to contribute to international EDA standardization groups, and to promote the use of EDA standards within the European community. These activities gained support from the Commission of the European Communities (CEC) in the frame of the ESPRIT program, and the first European CAD Integration Project was born.

ECIP, in two years, led all the organizations involved to the same understanding of the EDA standardization problems. Beginning in 1988, the ECIP project became fully mature with the direct involvement of eight major European microelectronics companies together with research and de-

0272-1732/92/0800-0054$03.00 © 1992 IEEE

velopment institutes in France, Germany, The Netherlands, Italy, and the UK, and the indirect involvement of many other European projects.

(The major companies are Bull, ICL (International Computer Limited, UK), Philips, SGS-Thomson, Siemens, Siemens-Nixdorf, Thomson-CSF, and later on Recal-Redac. The institutes are UCI Microelectronique, Les Ulis, France; University of Manchester; University of Paderborn; GMD (the German National Research Center for Computer Science); University of Hagen; Thomson-Sintra; IMT (Institut Mediterranean de Technologie, Marseille); and CNET (Centre National d'Etudes des Telecommunications, Paris.)

Such wide organization ensures the successful adoption of standards on a wide scale.

Some impact (adoption of interim recommendations) has already been achieved. ECIP has been a key contributor to the establishment of conceptual modeling techniques (information modeling) to underpin the consistency of EDIF and CFI (the CAD Frameworks Initiative).[5] ECIP has also stimulated interest in the 1992 ballot of the VHDL standard as well as in the evaluation of VHDL modeling practices and efficiency.

The key role of standards for the electronic industry was fully established when, in the beginning of 1991, the JESSI Subprogramme Management Application Board and the eight main companies involved in ECIP decided to found the European CAD Standardization Initiative. ECSI acts as the industrial steering committee of ECIP and holds responsibility for defining the strategy and goals of the project, validating the quality of its industrial results and promoting their use within and outside the European community. One key decision of this initiative was the introduction of technical centers or European test sites in three areas:

- Marseilles, France, and Munich, Germany, for VDHL;
- Manchester, UK, for EDIF; and
- Birlinghoven, Germany, for frameworks.

Other important initiative decisions are the establishment of a Permanent Standardization Office and, for the evaluation of the whole ECIP activities, the creation of an independent advisory board with worldwide-recognized experts. Figure 1 shows the organization of ECSI and ECIP.

EDA standardization activities

ECSI and ECIP currently support EDA standardization activities in the areas of languages, interfaces, and environments. But, the initiative is open to other EDA standardization activities.

VHDL. VHDL can specify a product without revealing design implementation characteristics or technological considerations. As a result VHDL presently forms the most promising contractual standard required to exchange design information, on a cooperative basis, either between industrial partners within electronic projects or between the customer and supplier of electronic products. Several applications are currently calling for VHDL-related exchanges of information concerning shared developments between industrial partners or within the scope of international contracts.

These first experiments however have demonstrated that—whatever the skill of the language designer—several unexpected weaknesses and a lack of critical performance exist in the present VHDL version. In fact, both design methodology expertise and enabling technologies are evolving concurrently.

To make VHDL practical for industrial applications, ECIP identified different activities and worked out the following:

1) *An active contribution to the standardization effort.* This contribution will address the ongoing restandardization of VHDL (to produce a new version by the end of 1992) and the standardization process of the VHDL companion standard: WAVES.
2) *A technical contribution to improve the use of the present VHDL version.* It includes modeling guidelines, package definitions, and modeling templates to support application domains such as formal proof, synthesis. The ex-

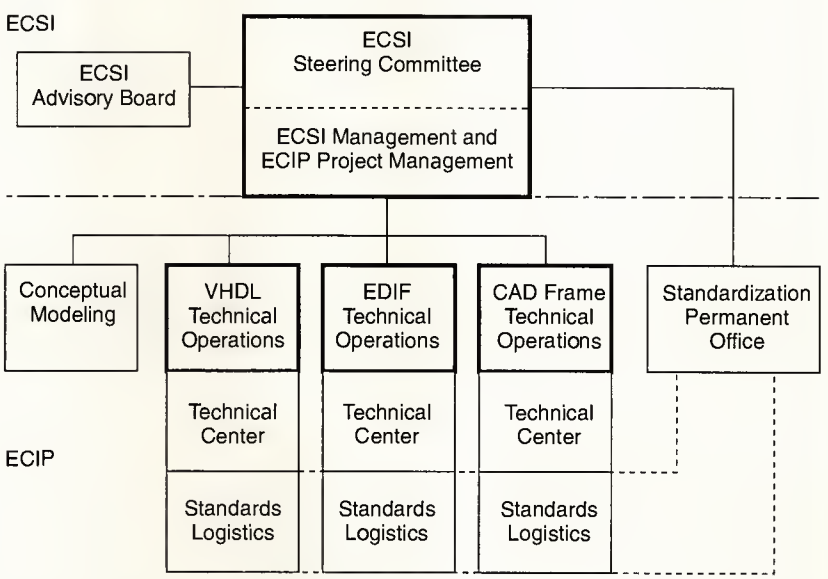


Figure 1. European CAD Standardization Initiative (ECSI) and ECIP project.

Glossary

CEC	Commission of the European Communities
CACE	Computer-aided concurrent engineering
CAD	Computer-aided design
CAE	Computer-aided engineering
CEN	European Committee for Standardization
CENELEC	Committee for Electrotechnical Standardization
CFI	CAD Framework Initiative
ECIP	European CAD Integration Project
ECSI	European CAD Standardization Initiative
EDA	Electronic design automation
EDIF	Electronic Data Interchange Format
EDIF TSF	EDIF test view proposal
ESPRIT	European Strategic Programme for Research in Information Technology
IEC	International Electrotechnical Commission
JESSI	Joint European Submicron Silicon Initiative
OMG	Object Management Group
OSF	Open Software Foundation
PCTE	Portable Common Tool Environment
VHDL	VHSIC Hardware Description Language
WAVES	Waveform and Vector Exchange Specification

tension of VHDL to describe analog behavior will be also an important topic in ECIP.

3) *Validation of the proposals concerning new features of the VHDL 1992 version.* The migration from the present version to the new version will be an important subject of work to provide VHDL users with practical and validated recommendations and appropriate work-around solutions.

Since 1990, the European involvement in VHDL has increased greatly. The Euro VHDL Conference, first organized in September 1990, convenes annually. In 1992, around 100 members from Europe voted on the VHDL 92 ballot on restandardization. The European Working Group on VHDL Standardization and the European Chapter of the VHDL Analysis and Standardization Group have met 10 or more times and produced 74 requirements for an improved VHDL standard.

The participation of European members in the VHDL 92 technical work process proved very important and should result in the coverage or consideration of more than 50 percent of the European requirements.

EDIF. This industry standard for transferring electronic design data particularly supports netlists and schematics. Furthermore it is intended to be flexible and extensible to meet the needs of a large group of different users throughout the world. Accordingly, most of the electronics industry, including EDA tool vendors, is already committed to EDIF, which is becoming a core component of EDA framework activities for data exchange. EDIF has been shown to be basically successful for the transfer of schematic and netlist data, and a number of enhancements to extend support in these areas are currently being developed.

Europe has already made a significant contribution to the adoption of EDIF as an industry standard. The results of trial transfers have been analyzed to determine the correctness and completeness of the intended data exchange. This analysis led to a series of recommendations being made on the use of EDIF for CAE/CAD data exchange. These recommendations are regularly updated by the ECIP team.

The following list of activities gives an indication of the European contribution in the area of EDIF development and support, mainly achieved through ECIP:

- The active participation of ECIP workers in the EDIF Technical Committee and a number of Technical Subcommittees provides a strong European influence on the development of EDIF.
- The expertise within Europe in information modeling provided tutorials and assistance to the EDIF community both in the United States and Europe. Thus EDIF developed detailed information models in the areas of connectivity, schematic capture, PCB, and test.
- Europe totally supports an EDIF questions-and-answers facility, which enables users to resolve technical queries regarding the use of EDIF.
- ECIP established a collection of test data files that is used to check the correctness of implemented interfaces.
- ECIP members actively participate in the organization of the annual EDIF forums and other EDIF-related awareness events, for example, tutorials.
- ECIP continues to develop EDIF validation and certification services that enable European users to qualify vendor (or internal) EDA tools with a high level of confidence.

Since EDIF is still an emerging standard, a number of important areas still need to be addressed: extension to PCB, test, VHDL interoperability, and validation activities. The EDIF Technical Committee has planned a program of work covering enhancements

- to EDIF schematics and basic enabling improvements,
- to include PCB representation, and
- for test.

EDIF actively worked out information models that are now available for all these areas at an earlier date than those proposed for the publication of the extensions. These models have been (or will soon be) transferred to the EDIF Technical Committee.

Generally speaking, ECIP works intensively in those areas of EDIF which, although incompletely specified, are the most industrially used or the most urgently needed. Taking into account its initial goal (providing an interchange mechanism capable of covering all aspects of system description), ECIP covers both VLSI (very large-scale integration) and PCB aspects. One of the areas needing a great deal of work is the support for PCB design data exchange.

Europe has played a major role in information modeling for PCB data. This activity will continue in strong relation with the EDIF Technical Committee. The evaluation through industrial trials of the new proposals for EDIF PCB view and other emerging view enhancements is vital also to the successful introduction of these enhancements.

Europe has also been particularly active in the testing domain. The ESPRIT Eureka project prepared a proposal for an EDIF test view (also called EDIF TSF), which is now promoted through ECIP in the EDIF Technical Committee.

Support of the work in Europe and the US on standard interoperability, in particular between EDIF and VHDL, is another key area requiring attention. The natural link here with the work being carried out on VHDL developments needs to be maintained. Mechanisms to link to other standards providing, for example, easy access to standards like IEEE library symbols also need to be investigated.

The work of the EDIF Technical Committee and Technical Subcommittees is of utmost importance to the success of EDIF. Europe maintains strong representation in the Technical Committee and in a number of Technical Subcommittees. This effort will continue to be used through ECIP to drive the EDIF standard and will be maintained at least at its present level. A beneficial side effect of this involvement is the early access of the industry to the proposed EDIF enhancements.

Frameworks. In 1989 the JESSI Subprogramme Application took over existing European framework developments in The Netherlands, France, Sweden, Italy, and Germany, and created a joint development project called JESSI CAD Frame. This project is the backbone of all CAD development in JESSI. Funded by ESPRIT (project 7364) since 1990, it is currently the only integration platform for all ESPRIT CAD projects.

The JESSI Common Framework, the new extended project name, is also the backbone for the EDA framework standardization activities in ECIP. These activities are combined in the Euro CFI organization under the umbrella of the European CAD Standardization Initiative. Euro CFI is the European partner of CFI (the CAD Framework Initiative, Inc. located in Austin, Texas[5]). The Europeans are committed to its mission: "Define interface standards that facilitate integration of design automation tools and design data for the benefit of end users and vendors worldwide."

CFI in the meantime counts a broad industry participation with over 50 members worldwide in the major semiconductor, systems, computer, and EDA companies. Corporate members in Europe include Siemens, Siemens-Nixdorf, Bull, Philips, SGS-Thomson, Alcatel, Genrad, Ericsson, and Racal-Redac.

Europe has played a major role in information modeling for PCB data and has been particularly active in the testing domain.

Focus areas for framework standardization are

- architecture,
- intertool communication,
- design representation,
- user interface,
- design methodology management,
- design data management,
- system environment,
- component information representation, and
- technology CAD.

Prior to February 1992, nine corresponding Technical Subcommittees—in which the European partners contributed generally through approved Technical Cochairs in each Technical Subcommittee—performed the work.

During the annual business meeting in February 1992, the CFI board of directors decided on a reorganization of the Technical Committees. The core groups are now the Planning Working Group and the Architecture Working Group. The work of the former Technical Subcommittees is now implemented in three new Technical Committees.

- *The Design Information Technical Committee* works on all development activities related to the electronic design domain (design representation, technology CAD, simulation services).
- *The Domain Independent Services Technical Committee* works on all development activities *not* related directly to electronic design and more generic in nature (intertool communication, design object management).
- *The Computing System Environment Services Technical Committee* works on low-level services necessary to support the framework environment (user interface, storage management, extension language).

The European partners have already nominated the European cochairs as representatives of the new organization and the CFI board of directors accepted this nomination.

Due to the strong involvement of the major computer companies (SNI, IBM, DEC, Sun Microsystems, Hewlett-Packard/Apollo), a framework based on the CFI standards will be developed and marketed as a software infrastructure just above the operating system level of hardware platforms. It will support the application environments of ECAD (electronic CAD), MCAD (mechanical CAD), and CASE (computer-aided software engineering).

Relationships with other standardization bodies like the Open Software Foundation are in development. PCTE and OMG are already well recognized as organizations, and contacts are already established.

The European Standardization Organization

Immediately after the foundation of ECSI the members recognized that a close link with the European Standardization Organization was essential to assure official approval of the results provided by ECSI. So, national representatives contacted CEN/CENELEC. CENELEC decided in December 1991 to start a new Technical Committee, TC 117 on electronic design automation. In its first meeting in April 1992 TC 117 incorporated the exact scope and goals of ECSI. In addition, the national members of CENELEC nominated ECSI members in TC 117 to build its core group. The elected chair of TC 117 also chairs the ECSI. Siemens holds the secretary position.

After the 1991 formation of ECSI, the foundation of the CENELEC TC 117 in 1992 is the important step toward achieving standards in the field of EDA in Europe. The main goals of ECSI and CENELEC are to contribute to the worldwide activities in EDA standardization. Both organizations should constitute the European focal points for the activities in the US with active contribution. Both organizations also support the usage of standards within the European microelectronics industry.

Major EDA exhibitions such as the Design Automation Conference and Euro DAC help demonstrate the European contribution and results. In addition, the European CAD Standardization Initiative remains open for additional activities such as libraries in the field of electronic CAD.

THE COMMISSION OF THE EUROPEAN COMMUNITIES actively supports the European EDA standardization activities via its ESPRIT program. Due to the founding of the European CAD Standardization Initiative by the major European companies, the former research-oriented activities gained important industrial promotion. Forming a new Technical Committee (TC 117) within CEN/CENELEC permitted the European EDA standardization activities to become the official entity of the European Committee for Standardization. Its activities are not limited to the current areas of interchange formats, description languages, and integration platforms; the structure is flexible and open to new groups.

Acknowledgments

ESPRIT project 2072 ECIP of the Commission of the European Communities partly supported this work. We thank all the participants in the ECIP project for their contributions

References

1. Richard Comerford, "Special Focus on Technology 1992: Software," *IEEE Spectrum*, Jan. 1992, pp. 30-32.
2. *Electronic Design Interface Format*, Version 2.0.0, ANSI/EIA 548, 1988; Version 2.0.101 Draft PN-3029, Design Automation Conf. Release, June 1992.
3. *IEEE Standard VHDL Language Reference Manual*, 1076-1987, (ANSI recognized), IEEE, Piscataway, N.J., 1987.
4. *IEEE Standard VHDL Language Reference Manual Interpretations*, 1076-Interpretations-1991, IEEE, 1991.
5. Andy Graham, "The CAD Framework Initiative Standards: Progress Toward First Publication at Year End," *IEEE DATC Newsletter on Design Automation*, Spring, 1992, pp. 13-21.

Anton Sauer serves as director of CAD Electronics and Engineering at Siemens Nixdorf Informationssysteme AG (SNI). He is currently responsible for SNI's Siframe development. He chairs the JESSI Subprogramme Management Application Board, the European CAD Standardization Initiative, and the new CENELEC Technical Committee on Electronic Design Automation.

Sauer received the Diploma for electrical engineers from the Technical University in Munich and a PhD in computer science from the Johannes-Kepler University in Linz, Austria. He is a member of The IEEE Computer Society and the German Gesellschaft fuer Informatik.

Jean Pierre Tual supervises the design and integration of software for board and VLSI layout and synthesis in the CAD department at Bull S.A. He also contributes to the management of numerous ESPRIT programs in the area of microelectronics. His main interests include VLSI layout and synthesis, CAD standards, and cryptography.

Tual received an engineering degree in computer science from the Ecole Nationale Superieure de Techniques Avancees in France and a PhD in mathematics from the University of Paris. He holds membership in the IEEE, SIAM American Mathematical Society, and Societe Mathematique de France.

Robin La Fontaine is an independent consultant and project manager for the ESPRIT ECIP. His interests include engineering data management, electronic data interchange, database design, and information modeling. He is a member of the EDIF Technical Committee.

La Fontaine read engineering science at Oxford University and holds a postgraduate degree in computer science. He is a Chartered Engineer and a member of the Institution of Mechanical Engineers in the UK.

Direct questions concerning this article to Anton Sauer, Siemens Nixdorf Informationssysteme AG, Dept. AP 41, PO Box 830951, D-8000 Munich 83, Germany; or e-mail at sauer@ztivax.siemens.com.

Floating-Point Processors Join Forces in Parallel Processing Architectures

A unique hardware architecture and extensive software support make the TMS320C40 floating-point digital signal processor an excellent parallel processing building block for high-performance applications. A particular advantage of the C40 comes from its six communication ports. We detail the hardware architecture and software capabilities of the C40 and explain, through examples, how it may be used in a parallel processing environment.

Ray Simar, Jr.

Peter Koeppen

Jerald Leach

Steve Marshall

Dave Francis

Greg Mekras

Jeffrey Rosenstrauch

Scott Anderson

Texas Instruments Incorporated

Conventional computers based on the serial machine architecture invented by John von Neumann half a century ago face formidable challenges in supporting modern applications. A serial machine's CPU data removal, manipulation, and return operations are too slow and inefficient to get the job done. An image processing system, for example, might be required to perform a series of 100 operations on millions of individual picture elements that represent the picture on a monitor. Using serial data techniques, a von Neumann machine would spend the vast majority of its time merely retrieving and storing picture data in central memory. It would leave virtually no time for computations necessary to alter screen images.

Rather than relying on one serial processor to handle all the operations on each element, we can distribute the operations on portions of the picture elements to a number of different processors. In addition to dedicating portions of the overall task to individual processors, all processors can execute their operations concurrently, depositing data in small local memory banks. Parallel processing is much better suited to the high-performance computing applications found in today's environment. The big difference between the two methods is that a parallel architecture devotes most of its time to processing data in parallel while a serial architecture bogs down when moving data in and out of storage.

Parallel computing concepts have been around for almost 25 years, but they've really come alive through the shrinking geometry of semiconductor technology. At VLSI dimensions, it becomes possible to fabricate processors small enough and with sufficiently low power dissipation to employ large numbers of them in conventionally packaged systems. Successful parallel computer designs contain from 32 to 64,000 processing elements.

Until now, parallel architectures used general-purpose microprocessors as the basic computing elements. But many high-performance applications can be handled even more efficiently using techniques first seen in digital signal processing. DSP chips have been a part of semiconductor technology for over 10 years, but now we've designed a special device specifically for parallel architectures.

Our DSP device is the TMS320C40 floating-point processor, which is among the highest performance 32-bit microprocessor devices available on the market. It operates at 275 million operations per second (MOPS) and transfers data at a rate of 320 Mbytes/s with a 40-ns cycle time.

A key architectural feature of the C40 for par-

0272-1732/92/0800-0060$03.00 © 1992 IEEE

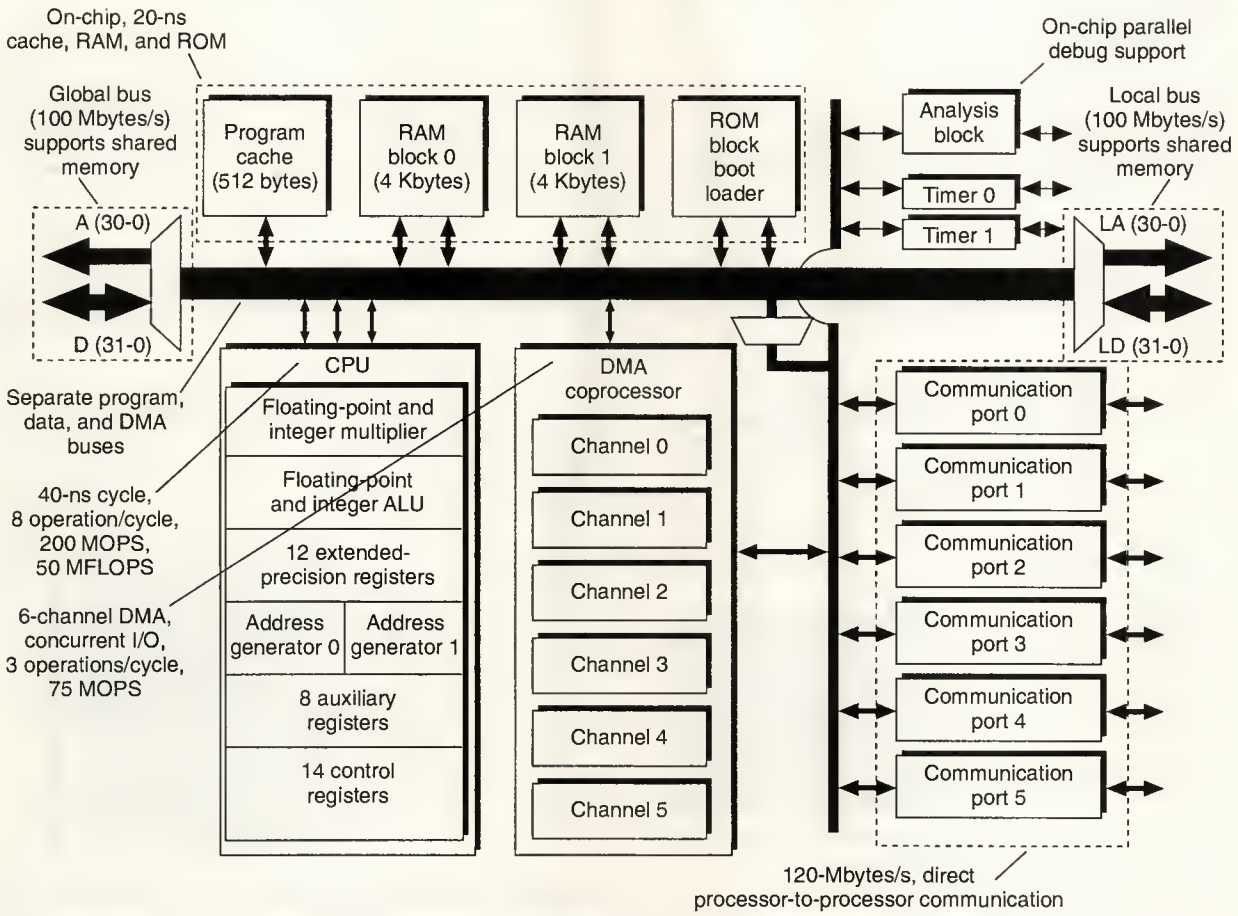


Figure 1. The TMS320C40 architecture.

allel computing is the six parallel bidirectional communication ports that permit direct connection and communication between processors in a parallel system. Any number of C40 processors can be linked together through the ports so a designer can custom tailor the size of a processing system and its level of performance. Moreover, the ports allow easy implementation of the critical automatic message buffering and synchronization that must take place between different tasks running on parallel processors.

The C40 features a unique on-chip DMA coprocessor to support interprocessor communications concurrently with calculations being executed in the CPU. This coprocessor supports six independent DMA channels (supporting each communications port) that provide concurrent I/O capabilities for parallel processing architectures and eliminate the I/O bottlenecks traditionally associated with such designs.

A full range of development tools designed for parallel processing accompanies the C40. Our ANSI-compatible, optimizing C compiler with a parallel processing runtime library supports code generation. SPOX, the first real-time operating system for embedded DSP systems (from Spectron Microsystems, Inc.), provides portable code to support the on-chip communication ports and DMA coprocessor. Other available state-of-the-art tools relieve designers of the time-consuming development tasks involved in parallel system programs.

Architectural features

The design team configured the C40 as a multitasking processor whose CPU is free and unencumbered to execute the algorithms vital to computationally demanding applications. They assigned interprocess communications to the communication ports and DMA coprocessor, while the global- and local-bus bandwidths (100 Mbytes/s each) are reserved for data and program fetches (Figure 1).

Capable of executing 200 MOPS at 40 ns, the CPU consists

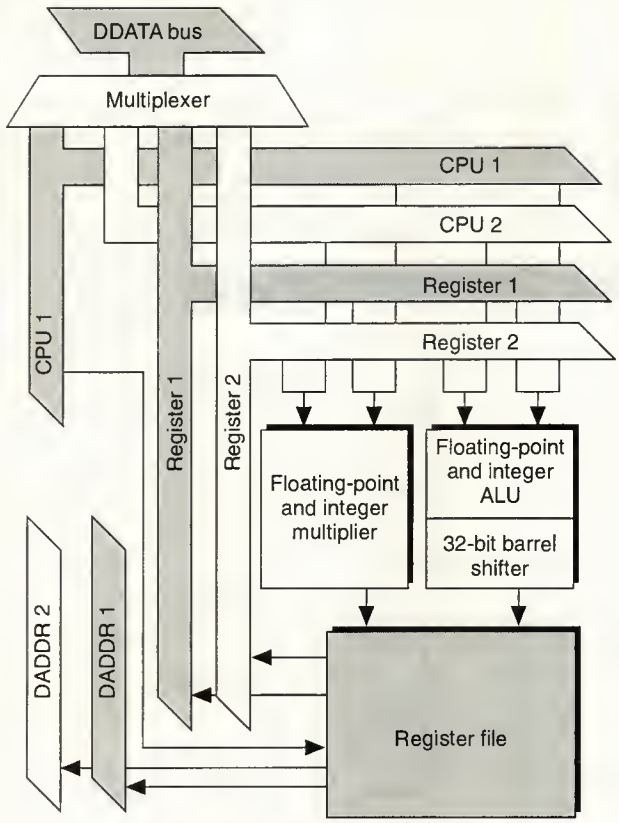


Figure 2. Multiplier and ALU operation.

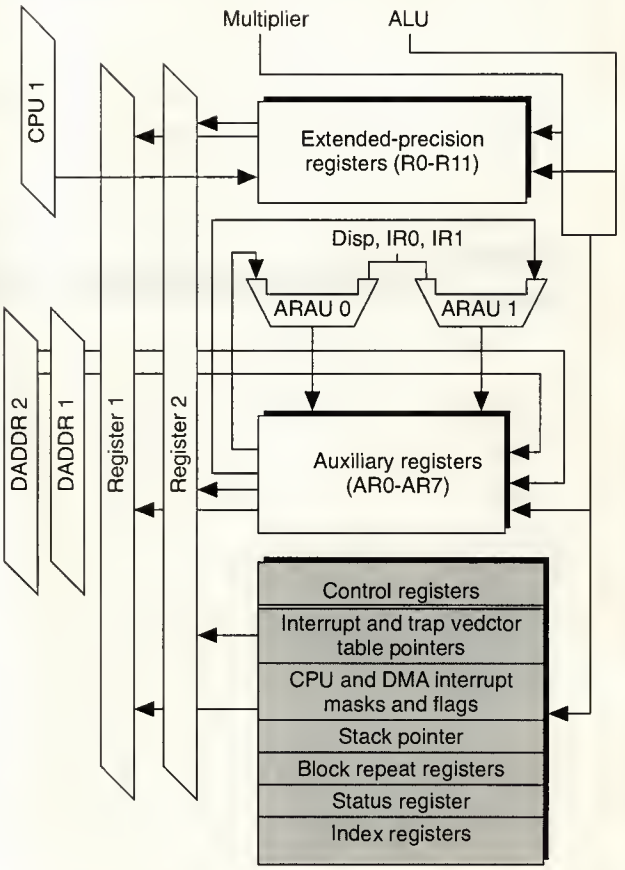


Figure 3. Register organization.

of the 40-bit floating-point/integer multiplier, ALU, 32-bit barrel shifter, 32-word primary register file, expansion register file, and two auxiliary register arithmetic units. Optimized for mathematically intensive applications, the CPU's architecture and instruction set use high-level languages to achieve high performance and small code size. On-chip hardware supports IEEE format conversion, division, and square root functions, and byte and half-word accessibility. The CPU is also source-code compatible with the TMS320C30 32-bit floating-point DSP.

The multiplier performs one-cycle, single-precision (32-bit) and extended-precision (40-bit) floating-point arithmetic operations as well as one-cycle, 32-bit integer multiplication. The ALU carries out one-cycle, single- and extended-precision floating-point arithmetic as well as 32-bit integer and 32-bit logical operations (Figure 2).

Twelve of the 32 registers in the primary register file are extended-precision units that support operations on single- and extended-precision floating-point numbers and 32-bit integers. The eight auxiliary registers are used primarily for addressing. Two auxiliary-register arithmetic units generate two addresses in one cycle and support addressing with immediate displacements, index registers (circular and modulo), and bit-reversals. The auxiliary registers also serve as 32-bit general-purpose registers. The remaining registers handle a variety of system functions. These include cache management addressing, stack management, processor status, zero overhead block repeats, interrupts, and relocatable interrupts and trap vectors (Figure 3).

For direct processor-to-processor communication, the C40 contains six asynchronous, high-speed communication ports, one of which is illustrated in Figure 4. Each port can transfer bidirectional data at 160 Mbits/s (5 Mwords/s), thus allowing high-speed interprocessor communication. In addition, each port independently buffers all input and output data transfers, provides automatic arbitration and handshaking for direct processor-to-processor connection, and supports synchronization between the CPU or DMA as well as among a wide variety of multiprocessor architectures. The 8-word, 32-bit input and output FIFOs can be read and written to

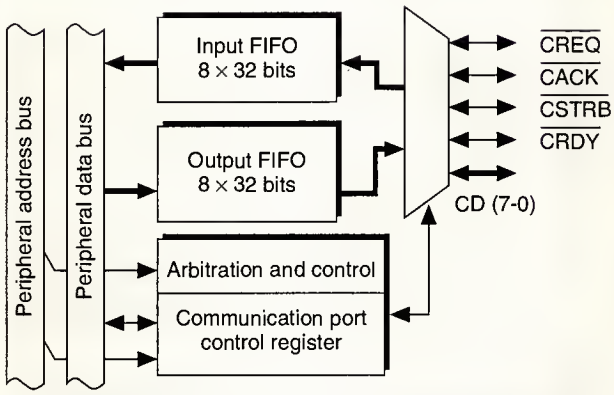


Figure 4. One communication port.

internally by either the CPU or the DMA coprocessor. FIFO inputs and outputs exist as 32-bit values in the processor's memory map.

The input FIFO buffers incoming data, and the output FIFO buffers outgoing data. Such buffering allows a high degree of decoupling between computations and communication overhead. Each FIFO is eight levels deep and 32 bits wide. When two C40 DSPs are connected via their communication ports, the effective FIFO depth is 16 levels.

Within each communication port is a PAU, or port arbitration unit, that handles data movements to and from the port. When any number of C40s are interconnected, the PAUs cooperate in generating the signals and control sequences that permit data transfers at the highest possible rate. Since communication ports are bidirectional, PAUs arbitrate for ownership of the communication-port data bus as well as effect alternate ownership. The communication ports can be connected between processors without glue logic.

The DMA coprocessor (which can autoinitialize) supports six DMA channels that execute fast data transfers to and from anywhere in the C40's memory map, which includes the communication ports. A special split mode can support 12 channels for dedicated communication ports to and from memory transfers. CPU and DMA operations execute concurrently, which maximizes the chip's processing power and provides great throughput.

Source and destination address registers within the coprocessor (Figure 5), with variable indices, permit stepping through matrices either by row or by column. Link pointer registers allow the DMA channels to reprogram themselves between jobs, thus maximizing sustained CPU performance by not requiring the CPU to reprogram each channel after each transfer. Support is also provided for the block transfer of data with bit-reversed addresses (useful in fast Fourier transform applications). External and internal interrupts synchronize data transfers.

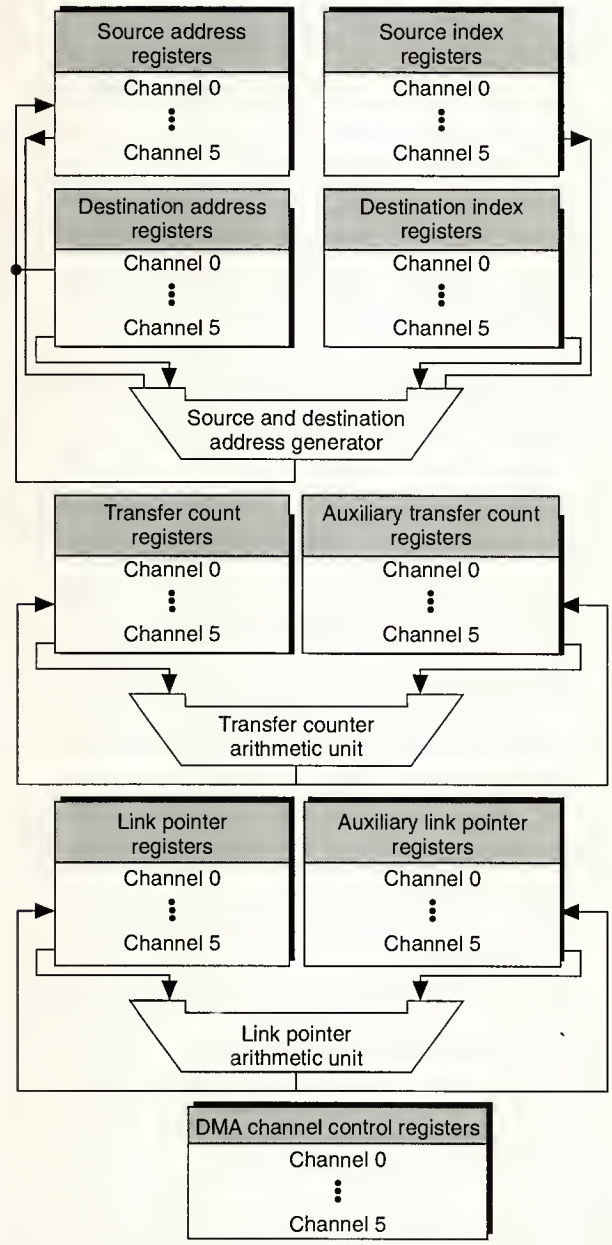


Figure 5. C40 on-chip DMA coprocessor.

Whenever a block transfer completes, the DMA can be programmed to execute several tasks. It can signal that the transfer is complete, stop the DMA channel until reprogrammed, and automatically reinitialize its registers via linked lists stored in memory. This autoinitialization capability allows the DMA to run continuously without CPU intervention.

The C40 boasts a total memory space of 16 Gbytes (4 bil-

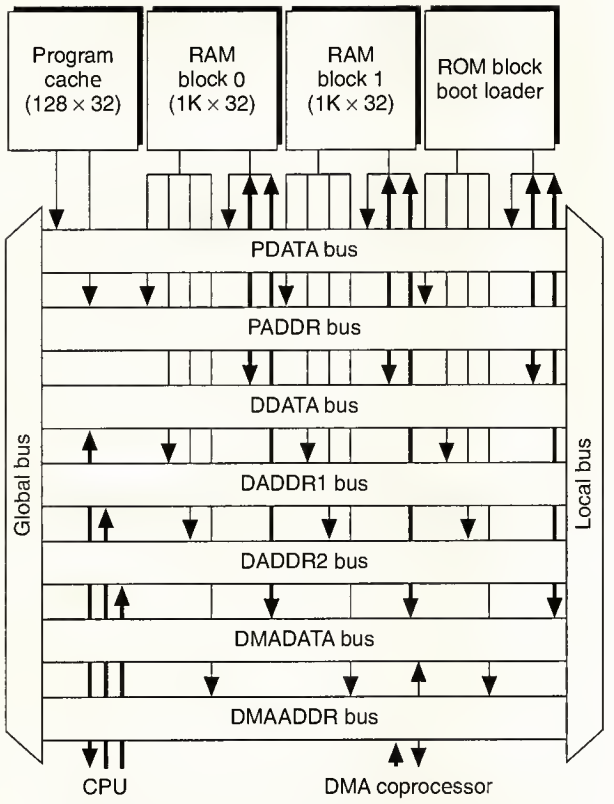


Figure 6. Memory organization.

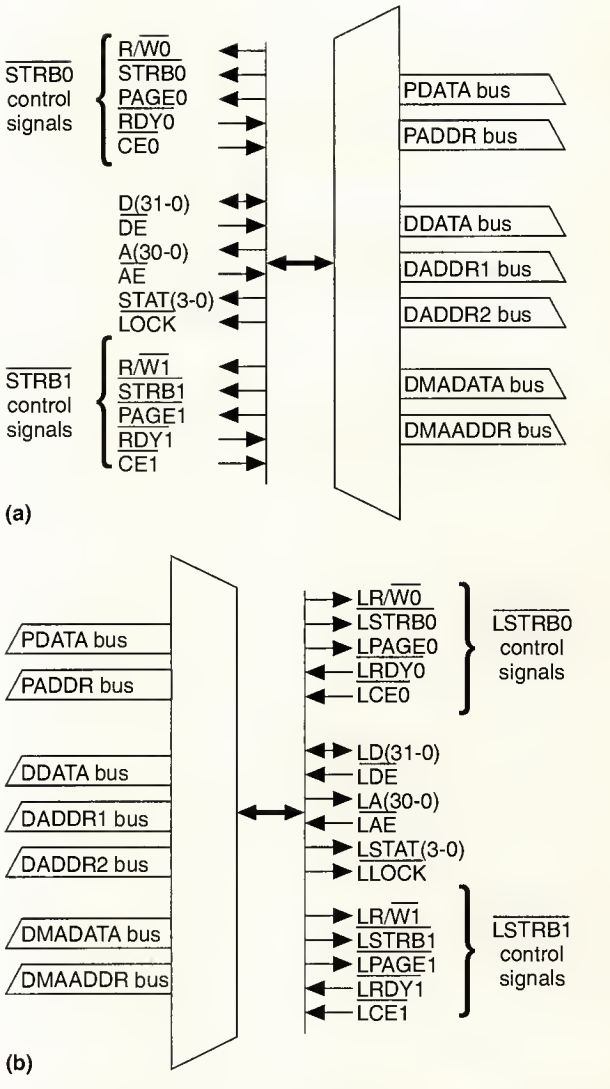


Figure 7. Global (a) and local (b) memory interfaces.

lion 32-bit words) and provides an 8-Kbyte RAM and a 16-Kbyte ROM on chip (Figure 6). Each RAM block and ROM block can support two data accesses in one cycle. Contributing to maximum performance at a minimum cost is a 512-byte, on-chip program cache that stores often-repeated sections of code to reduce the number of off-chip accesses the DSP must make. Less frequently used code can be stored off chip in slower and lower cost memories.

Key to the C40's high performance are two identical, 80-pin parallel interfaces known as the global and local memory interfaces (Figure 7a,b). These asynchronous buses can be placed in a high-impedance state for multiprocessor bus sharing in parallel systems. Each interface supports the large number of separate internal program, data, and DMA buses. Both interfaces consist of separate 32-bit data and 31-bit address buses and operate via independent enable signals for data, address, and control. They incorporate look-ahead bus status signals for defining current and requested bus operations to support intelligent bus arbitration.

A unique feature of the C40 is its dedicated, on-chip analysis module for debugging problems associated with parallel processing (Figure 8). This module contains a test port compatible with the JTAG/IEEE 1149.1 standard for developing and testing hardware.[1] This interface, along with some special-purpose signals, allow the debugging of any number of C40 DSP chips connected in a parallel system. Other elements of the analysis module and chip are breakpoint comparators for program, data and DMA accesses, a program trace stack, and an event counter for benchmarking.

An important feature of the module is the use of the JTAG port for communication with TI's XDS510 in-circuit emulator. The XDS510 allows any number of C40s in a parallel system

READER SERVICE CARD

IEEE MICRO

August 1992 issue (card void after February 1993)

Name _____

Title _____

Company _____

Address _____

City _____ State _____ Zip _____

Country _____ Phone(_____) _____

Reader Interest

(Add comments on the back)

Readers,
Indicate your interest in articles and departments **by circling the appropriate number** (shown on the last page of articles and departments):

150	151	152	180	181	182
153	154	155	183	184	185
156	157	158	186	187	188
159	160	161	189	190	191
162	163	164	192	193	194
165	166	167	195	196	197
168	169	170	198	199	200
171	172	173	201	202	203
174	175	176	204	205	206
177	178	179	207	208	209

Product Information

(Circle the numbers to receive product information)

1	21	41	61	81	101	121	141
2	22	42	62	82	102	122	142
3	23	43	63	83	103	123	143
4	24	44	64	84	104	124	144
5	25	45	65	85	105	125	145
6	26	46	66	86	106	126	146
7	27	47	67	87	107	127	147
8	28	48	68	88	108	128	148
9	29	49	69	89	109	129	149
10	30	50	70	90	110	130	
11	31	51	71	91	111	131	
12	32	52	72	92	112	132	
13	33	53	73	93	113	133	
14	34	54	74	94	114	134	
15	35	55	75	95	115	135	
16	36	56	76	96	116	136	
17	37	57	77	97	117	137	
18	38	58	78	98	118	138	
19	39	59	79	99	119	139	
20	40	60	80	100	120	140	

READER SERVICE CARD

IEEE MICRO

August 1992 issue (card void after February 1993)

Name _____

Title _____

Company _____

Address _____

City _____ State _____ Zip _____

Country _____ Phone(_____) _____

Reader Interest

(Add comments on the back)

Readers,
Indicate your interest in articles and departments **by circling the appropriate number** (shown on the last page of articles and departments):

150	151	152	180	181	182
153	154	155	183	184	185
156	157	158	186	187	188
159	160	161	189	190	191
162	163	164	192	193	194
165	166	167	195	196	197
168	169	170	198	199	200
171	172	173	201	202	203
174	175	176	204	205	206
177	178	179	207	208	209

Product Information

(Circle the numbers to receive product information)

1	21	41	61	81	101	121	141
2	22	42	62	82	102	122	142
3	23	43	63	83	103	123	143
4	24	44	64	84	104	124	144
5	25	45	65	85	105	125	145
6	26	46	66	86	106	126	146
7	27	47	67	87	107	127	147
8	28	48	68	88	108	128	148
9	29	49	69	89	109	129	149
10	30	50	70	90	110	130	
11	31	51	71	91	111	131	
12	32	52	72	92	112	132	
13	33	53	73	93	113	133	
14	34	54	74	94	114	134	
15	35	55	75	95	115	135	
16	36	56	76	96	116	136	
17	37	57	77	97	117	137	
18	38	58	78	98	118	138	
19	39	59	79	99	119	139	
20	40	60	80	100	120	140	

SUBSCRIPTION CARD

SUBSCRIBE TO IEEE MICRO

All the facts about today's chips and systems

❑ **YES**, sign me up!

If you are a member of the Computer Society or any other IEEE society, pay the member rate of only $23 for a year's subscription (six issues).

Society: _____

IEEE membership no: _____

Society members: Subscriptions are annualized. For orders submitted March through August, pay half the full-year rate ($11.50) for three bimonthly issues.

Full Signature _____ Date _____

Name _____

Street _____

City _____

State/Country _____ ZIP/Postal Code _____

❑ **YES**, sign me up!

If you are a member of ACM, ACS, BCS, IEE (UK), IEEE but not a member of an IEEE society), IECEJ, IPSJ, NSPE, SCS, or other professional society, pay the sister-society rate of only $39 for a year's subscription (six issues).

Organization: _____ Membership no: _____

For information on airmail option, send fax request to (714) 821-4010.

❑ Payment enclosed *Residents of CA, DC, Canada, and Belgium add applicable sales tax.*

❑ Charge to ❑ Visa ❑ MasterCard ❑ American Express

Charge-card number

_____

Expiration date

____ ____
Month *Year*

Prices valid through 12/31/92

0892 MICRO

Charge orders also taken by phone:
(714) 821-8380 8 a.m. to 5 p.m. Pacific time
Circulation Dept.
10662 Los Vaqueros Circ., PO Box 3014
Los Alamitos, CA 90720-1264

Editorial comments

I liked: _____

_____

_____

I disliked: _____

_____

_____

I would like to see: _____

_____

_____

Reviewers Needed. *If interested, send professional data to Dante Del Corso, Dipartimento di Elettronica, Politecnico di Torino, C.so Duca degli Abruzzi, 24, 10129, Torino, Italy.*

For reader service inquiries, see other side.

PLACE POSTAGE HERE

PO Box is for reader service cards only.

IEEE Micro
PO BOX 16508
NORTH HOLLYWOOD CA 91615-6508
USA

Editorial comments

I liked: _____

_____

_____

I disliked: _____

_____

_____

I would like to see: _____

_____

_____

Reviewers Needed. *If interested, send professional data to Dante Del Corso, Dipartimento di Elettronica, Politecnico di Torino, C.so Duca degli Abruzzi, 24, 10129, Torino, Italy.*

For reader service inquiries, see other side.

PLACE POSTAGE HERE

PO Box is for reader service cards only.

IEEE Micro
PO BOX 16508
NORTH HOLLYWOOD CA 91615-6508
USA

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO. 38 LOS ALAMITOS, CA

POSTAGE WILL BE PAID BY ADDRESSEE

IEEE COMPUTER SOCIETY
PO BOX 3014
LOS ALAMITOS CA 90720-9804
USA

to be single-stepped, either individually or simultaneously, or halted and started independently or simultaneously. A breakpoint on one processor can halt any or all processors in a parallel system. And any single processor can be debugged without disturbing the operation of other processors in the system.

The C40's event counter serves a variety of purposes in system operations. It can count machine clocks to permit accurate profiling important to embedded design applications. It also can count operations such as instruction acquisitions, subroutine calls and returns, interrupts and traps, and breakpoints.

The software side

A full assortment of development tools is available to support C40 parallel processing applications. The tools focus on three areas: code generation, simulation, and hardware development and verification. Designed with a large, flexible register file, a software stack, and a large, continuous memory space, the C40 implements a high-level language compiler. A C compiler increased the transportability of applications that were tested on large, general-purpose computers, and decreased their porting time.

The ANSI-compatible, optimizing C compiler with a parallel processing runtime library supports code generation. The compiler is one of the first tools designed specifically for real-time, embedded applications, and the library simplifies the development of algorithms for parallel processing by supporting key operations such as message sending and receiving.

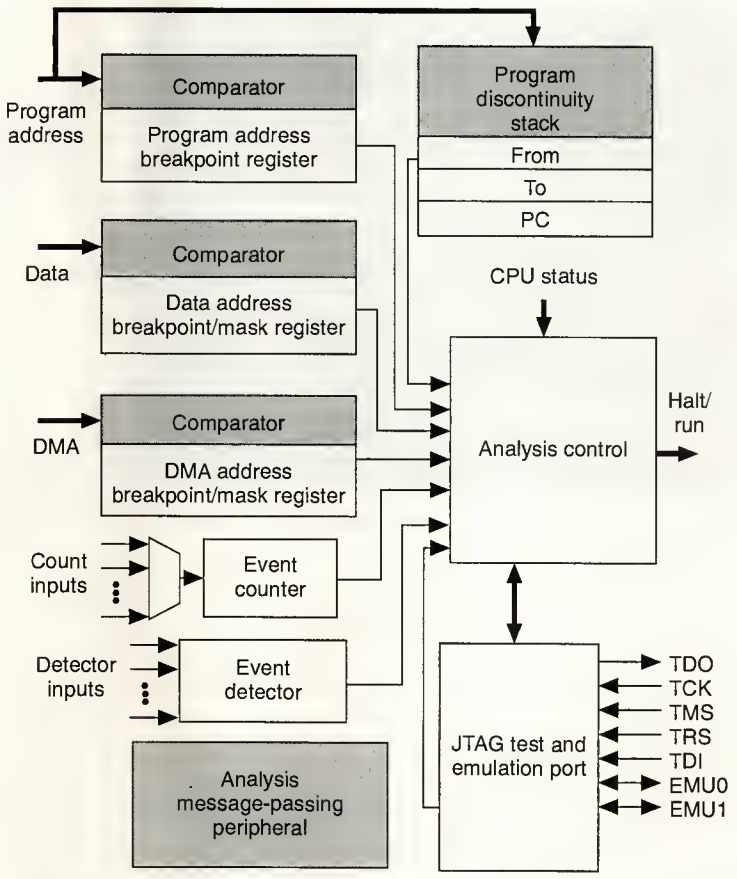


Figure 8. On-chip analysis module.

The compiler performs global and loop optimization such as strength reduction and analyzes code to optimize the use of memory and register variables. It also searches out vector and matrix operations and maps memory access to those C40 addressing modes that are optimized for vector and matrix operations. Programming can be accomplished through either the C compiler or via assembly language. Tartan, Inc. offers an Ada compiler with similar optimizing capability.

The SPOX real-time operating system for embedded systems provides drivers for the C40's communication ports and DMA coprocessor. It also provides a complete set of parallel processing primitives (Send Message, Receive Message, and so on) and offers a rich core of software functions for parallel processor applications.

SPOX covers a complete set of signal-processing functions for memory allocation, stream I/O, and DSP math operations. It allows C40 application programs distributed across many processors to trade data through their communication ports. A real-time, multitasking kernel permits the C40 to manage a number of complex tasks simultaneously.

Simulation and benchmarking are achieved through a variety of software tools. TI's state-accurate simulator with a high-level language debugger is a vehicle for the analysis of a C40. It allows detailed timing analyses of the critical portions of C and assembly-language programs.

Logic Automation, Inc. offers software behavioral models of the C40. This set of hardware verification models lets system designers simulate several C40s hooked together with their associated memory subsystems for analysis and evaluation of total system performance. The behavioral models can run on hardware platforms from Mentor Graphics, Valid Systems, and others.

Streamlined emulation of hardware development and verification gives systems engineers parallel processing debug capability and control over an unlimited number of parallel processors. TI's XDS510 in-circuit emulator, which works closely with the C40's on-chip analysis module, provides these

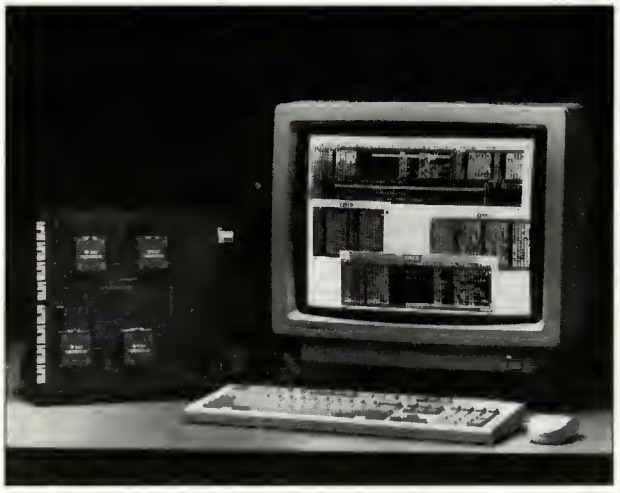


Figure 9. XDS510 in-circuit emulator.

capabilities (Figure 9). The embedded analysis function monitors overall operations of the C40 and provides designers with a visual report card of processor activity.

Parallel processing

Parallel processing is becoming increasingly important for a number of reasons. The first is simply the power of a parallel processor architecture compared to that of a single processor. When high-speed data manipulation is necessary in an application such as imaging, neural networks, or array processing, parallel devices can bring more computing power to bear. Fortunately, the declining cost of TMS320 chips is making parallel processing a cost-effective reality for many applications.

Parallel processing also shows promise for traditional DSP applications that formerly relied on one processor, since classical DSP algorithms are inherently suited for task partitioning. The tasks can be assigned to multiple processors. Examples of these classical problems are filtering, correlation, and FFTs, each of which can be represented by a signal flow graph. These graphs, which identify lower level functions by their parallel interactions, by their very nature imply suitability for parallel processing techniques.

C40 devices can serve as the foundation or building blocks for virtually any kind of parallel processor architecture needed to solve computationally intensive problems. In the most traditional approach, designers attach a global memory to the global bus of each of a number of C40 processors in a system so the processors can share common data (see Figure 10a). At the same time, each processor connects to a local memory over its local bus for storing private data. Such a parallel configuration could be applied to improve the performance of computer

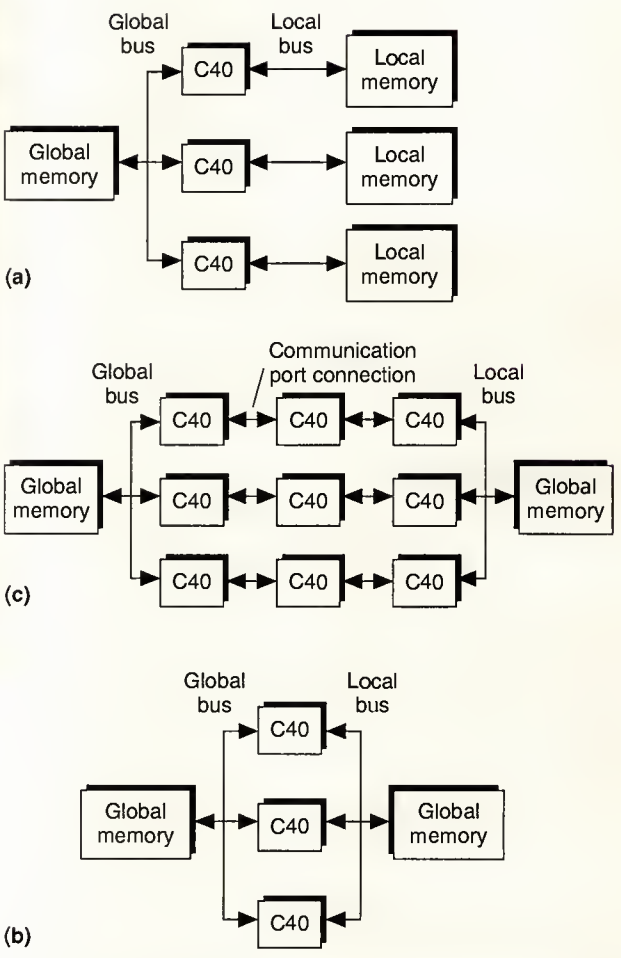


Figure 10. C40 parallel DSP system architectures: shared global and private local memories (a), shared global memory on the global and local buses (b), and shared memory and communication ports (c).

bus architectures such as the Nubus, Futurebus+, and VMEbus.

A variation on the above theme is a scheme in which global memory interfaces to both the shared and local buses (Figure 10b). This technique works with low-cost, small form-factor graphics applications involving perspective transforms and lighting. Figure 10c illustrates a parallel system in which C40 processors interface to each other via their communication ports, while both the local and shared buses are hooked to a global memory.

High-performance, pipelined linear-array processors using systolic arrays are another application of C40 DSPs (Figure 11a). Such an architecture is useful in the convolution, correlation, and complex arithmetic found in radar system appli-

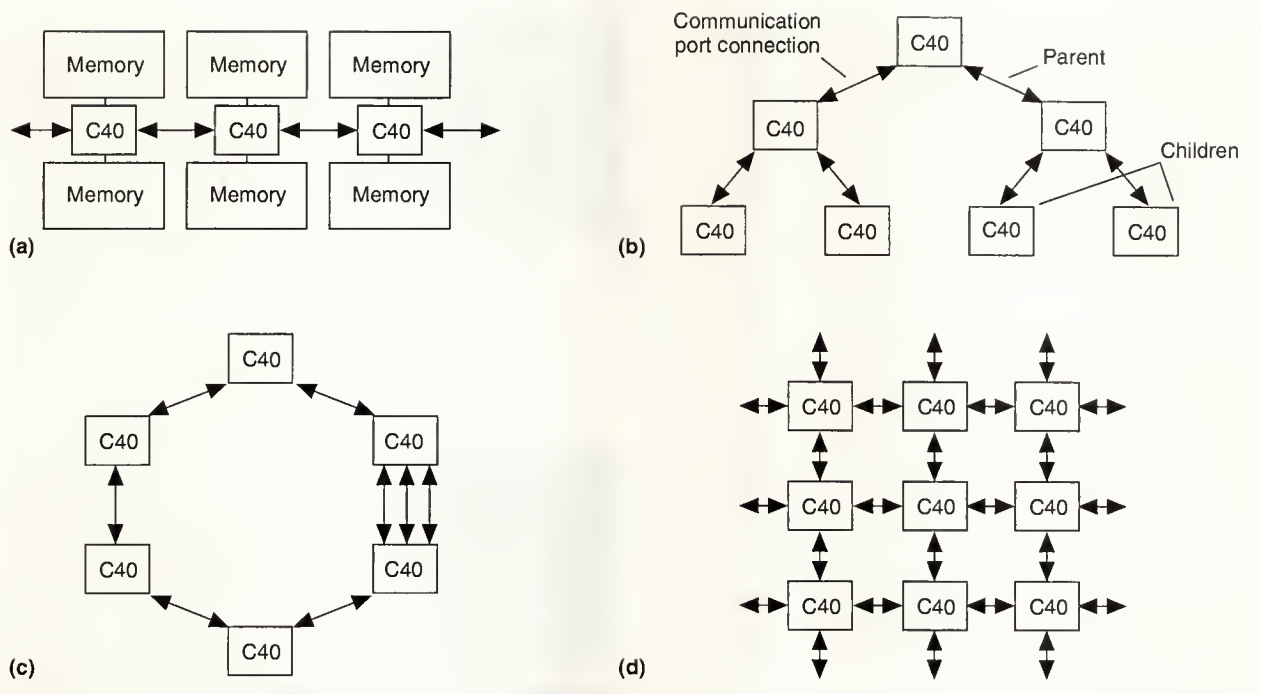


Figure 11. C40 system architectures: pipelined linear array (a), tree structures (b), clockwise and counterclockwise data flow (c), and 2D array (d).

cations. Tree structures (Figure 11b) allow rapid fan-out of data from "parent" processors to "children" processors to enable parallel processing. This efficient technique helps solve problems concerning speech recognition, database management, and encryption. The bidirectional ring (Figure 11c) with its clockwise and counter-clockwise data flow possibilities and C40 floating-point multiply and accumulate functions works well in machine learning and other applications that require neural networks. Connecting more than one communication port between two C40s increases communications bandwidth. The two-dimensional, hexagonal array of Figure 11d can provide the necessary matrix and vector performance required in applications such as numerical analysis and control theory.

Imaging systems have requirements that demand high performance, high bandwidth, large amounts of data, and real-time processing, needs that are supplied by multiprocessing and the C40. Kodak Research Labs has, for example, developed an image processing system for internal use based on the C40 that attaches to a Sun Sparcstation through one Sbus expansion slot. The system hardware consists of an Sbus card with one C40 with 32 Kwords of SRAM and 1 Mword of page-mode DRAM connected to an expansion board. The board contains six C40s, each with 32 Kwords of SRAM and 1 or 4 Mwords of page-mode DRAM. The C40s are configured in a fully connected network using the built-in communication ports. The expansion board connects to the single C40 Sbus board through communication port connectors on the Sbus backplane and allows C40s to be used in parallel to perform image processing tasks.

Another good example of C40 capabilities can be seen in work conducted at Accurate Automation Corporation. This firm has recently won a contract from the US National Aeronautics and Space Administration to adapt neural networks based on the C40 for use in the Space Shuttle robotic arm, also known as the remote manipulator system. The RMS has six degrees of freedom and can manipulate and control objects in space while being controlled by an astronaut inside the shuttle.

The current design of the arm gives operators trouble in performing some tasks. Accurate Automation plans to use neural networks as a way to ease the use and control of the system. The neural network of C40s allows the RMS to plan, coordinate, and smooth the coordination of the arm. The system also allows astronauts to use the arm in two ways. Assistant mode lets the network correct and control the arm as it is being guided by an operator, and autonomous mode lets the operator specify a point that the arm directs itself to automatically. The neural network's ability to learn from experience allows the arm to learn the relationships between movement and joint angle trajectories.

AS TODAY'S PERFORMANCE-HUNGRY SYSTEMS DEMAND more computing power than any single processor can deliver, parallel processing systems will increasingly become a common means of cost effectively achieving the designer's performance goals. For chip designers, the road ahead is marked by hardware and software design challenges. These include methods to simplify interprocessor communication, development tools to support parallel processing, and in-circuit emulation for debugging parallel connected processors and others. The TMS320C40 will support this new generation of devices devoted to parallel computing.

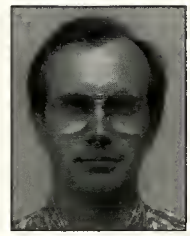
Parallel processing is becoming the most viable technique to deal with high-performance applications such as image processing, three-dimensional graphics, video conferencing, neural networks, and others. These applications are beyond the capabilities of a single DSP device, despite the fact that the performance of these chips has increased dramatically in the past few years. Even as performance increases however, device costs continue to fall. This makes it possible for designers to incorporate multiple C40 devices in a parallel processing system to obtain a cost-effective solution to a range of applications that demand very high performance.

References

1. *IEEE Standard Test Access Port and Boundary-Scan Architecture*, 1149.1-1990, (ANSI recognized), IEEE, Piscataway, NJ 08855-1331.
2. *TMS320C4X User's Guide*, Texas Instruments Incorporated, 1991.
3. *TMS320C4X Technical Brief*, Texas Instruments Incorporated, 1991.

Ray Simar, Jr., a senior member of the technical staff and advanced DSP architecture manager at Texas Instruments Semiconductor Group, is responsible for the development of advanced floating-point DSPs. He is the principle architect of the TMS320C40 parallel processor and the TMS320C30 floating-point, 32-bit DSP. He was the program manager for these products during their development, coordinating the design, development tools, and marketing activities.

Simar graduated with a BS degree from Texas A&M and holds an MS degree in electrical engineering from Rice University. He is a member the Institute of Electrical and Electronics Engineers and the Association of Computing Machinery.



Peter Koeppen is currently design manager of floating-point DSPs at TI. His technical interests include design methods and DSP algorithms. Previously, he worked with the TI LAN chip and 4-bit processor designs.

Koeppen holds BSEE and MSCS degrees from the University of Wisconsin-Milwaukee and is a member of the IEEE.

Jerald Leach is a senior member of the technical staff at TI and was the lead designer for the C320 DSP and C40 parallel DSP. Previous design experience includes general-purpose and graphics microprocessors and calculators.

Leach received the BSEE degree from Texas A&M.

Steve Marshall is currently a member of the group technical staff in application-specific design at TI, Houston. He has worked on logic and circuit design for several members of the TMS320 DSP family.

Marshall received a BSEE degree from the University of Cincinnati and is a member of the IEEE.

David E. Francis is currently a design engineer in TI's Application-Specific Products Group. He has been involved in full-chip logic simulation and fault grading.

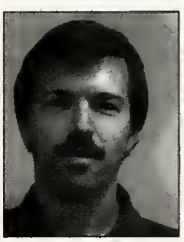
Francis received his BSEE degree from the University of Cincinnati and is a member of the IEEE.

Greg Mekras is assigned as a senior design engineer in TI's Customizable Digital Signal Processing Group. Previously, he served as physical layout design manager for the C40 and design manager for the TMS320E14 control-DSP EPROM microprocessor and design/product engineer on other TMS320C1X devices.

Mekras holds a BS degree in engineering physics from Murray State University, Murray, Kentucky.

Jeffrey Rosenstrauch is currently a design engineer in the same Application-Specific Products group, where he has been involved in automatic test pattern generation for fault grading.

Rosenstrauch received his BSEE degree from San Diego State University, California.



Scott Anderson is the primary product engineer with the C40 and worked as a design engineer with behavioral model development and logic simulation for the C40. He has also worked as a product engineer on a variety of members of the TMS320 family.

Anderson holds a BS degree in electrical engineering from Bradley University in Peoria, Illinois.

Direct any questions about this article to Ray Simar, Jr., Texas Instruments Incorporated, 12203 S.W. Freeway, Stafford, TX 77477.

Reader Interest Survey

Indicate your interest in this article by circling the appropriate number on the Reader Service Card.

Low 168 Medium 169 High 170

Special Report:
MITI's Real World Computing Program

A 10-year program will lay the technological foundations for the next century's information society. MITI expects future information systems to be based on a flexible integration of massively parallel computing, optical computing, neural computing, and logic programming. The Real World Computing program aims to establish theoretical foundations for these technologies, explore applications, and study how they can be integrated.

David Kahaner

US Office of Naval Research

[David Kahaner is on assignment with the US Office of Naval Research. He generally comments on activities in the Far East for inclusion in the Software Report column. Since we felt readers would be interested in a detailed description of Japan's 10-year computing plans, we also offer this special report. His comments are his own; they do not express any official policy.—Ed.]

The Real World Computing program (also called the New Information Processing Technology program) is a 10-year program sponsored by Japan's Ministry of International Trade and Industry (MITI) to lay the technological foundations for the information society that Japan anticipates in the next century. The RWC program replaces the well-known and much-discussed Fifth Generation project. This report summarizes the draft prepared by the RWC program's Feasibility Study Committee as presented in Tokyo March 2-3, 1992, at what was called RWC 92, the Second NIPT Workshop. My earlier reports in the June 1991 (p. 11) and June 1992 (p. 71) issues of *IEEE Micro* also dealt with this program.

MITI expects information systems of the next century to be based on several technologies (as opposed to a single, key technology): massively parallel computing, optical computing, neural computing, and logic programming. The RWC program aims to establish theoretical foundations for these technologies, explore applications, and study how they can be integrated. Flexible integration is an important goal for information systems that can deal with real-world problems. Program organizers see computing systems as evolving from conventional computing with numbers, documents, and data through fifth-generation (logic-based) computing, in which knowledge is processed, to flexible information processing, in which intuitive information can be processed.

Flexible processing is the heart of the RWC program. Example applications include incompletely specified (ill-defined) problems such as understanding situations in a noisy environment, large-scale problems such as simulating social and economic phenomena, and real-time problems such as constructing man-machine interfaces with virtual reality.

The RWC program professes two main technological aims:

- *To develop the computational bases.* This effort includes separate research on general-purpose massively parallel systems and special-purpose neural systems, optical computing and optical devices, and system integration.
- *To formulate theoretical foundations and then develop "novel" functions.* Theoretical foundations include all aspects of representation, storage, and recall of information;

0272-1732/92/0800-0070$03.00 © 1992 IEEE

information integration and evaluation; and learning and self-organization. Novel functions include flexible recognition and understanding of multimodal information, flexible inference and problem solving on flexible information bases, flexible or autonomous control, and flexible interactive environments for man-machine interfaces.

The RWC program plans a research partnership that will actually begin activities in October 1992 with a call for participation and subcontractors. Foreign firms and individuals can participate in the program by paying an initiation fee and joining the RWC Partnership. Since the application deadline will be near the end of 1992, substantial money will not enter the program until late 1992; the 1992 fiscal year budget is only about 900 million yen.

The workshop I report on here was closed to all but participants of earlier workshops and a few others who made special requests to the organizers. Non-Japanese participants came from Korea, France, the United Kingdom, Canada, Germany, The Netherlands, Italy, Australia, Singapore, and the United States. I was the only US attendee, although M. Miyahara from the NSF office at the US embassy also attended. Shun-ichi Amari of the University of Tokyo chaired the workshop steering committee and Toshitsugu Yuba of ETL served as vice-chair. Approximately 100 people attended the workshop.

Minor changes in the plan

Some minor changes in the RWC program made at this second workshop include the merging of the previously separate theory and novel-functions subprograms. Brain research has been deemphasized. Splitting the program into two five-year parts to allow a culling of less productive aspects does seem new—or newly emphasized. Also new is the proposal to build several prototype systems during this first five-year period, including one "platform" system for software development. But I do not understand how researchers are going to start immediately doing the theoretical and novel-functions research on systems that are not yet built.

The neural and parallel systems are now viewed as distinct. The parallel system will be on the order of one million processors in about 1,000 modules, and the neural system will be about one million neurons on about 100 wafers.

The list of research topics has been defined somewhat more carefully, but there are still far too many to be supported by even the most generous program—there are enough research topics listed to keep an army of researchers busy for decades.

MITI also changed or clarified its policy on who can be a member of the RWC Partnership. It now allows participation by organizations other than the original workshop participants. Partners will pay 10 to 50 million yen to join. While this is not much for companies, it is a lot for universities. Organizations can avoid partnership charges by working as subcontractors, but then royalties are much reduced.

I assume that many Japanese companies will want to become members for prestige and to be "plugged in," as much as for any other reason. The same might be said about some countries' motives for participating. ETL is generally viewed as the big "winner" in research funding. The Japanese government law changing the distribution of intellectual property rights does *not* apply to software; MITI staff admit that this is a prob-

Fundamental policies of the RWC program

The RWC program will be organized around the following fundamental policies:

1) **A central laboratory** (probably at or near the government's Electrotechnical Laboratory (ETL) in Tsukuba). Here common research will be performed, and students and postdoctoral researchers trained. For more highly individual research, several distributed laboratories will be established (perhaps not all in Japan).

2) **A competitive principle.** The first half of the program will support a large number of different approaches. For the second half, research themes will be evaluated and concentrated. Workshop steering committee vice-chair Toshitsugu Yuba explains that the program deliberately has a clear description but only vague, multiple subtargets.

3) **Interdisciplinary and international cooperation.** The program will actively pursue joint research with ETL, universities, and so on, and subcontract research from Japanese and foreign research organizations. According to Yuba, MITI sees this program as a real change of direction for its support of science, from one that invests in near-market development to one that is fundamentally research oriented.

4) **Publication of research achievements** to promote true international cooperation and open reporting via conferences and symposia. As much as possible, the RWC program will document results in English. However, many informal and internal documents will be in Japanese.

5) **Establishment of an infrastructure for research.** A high-speed network linking more than 10 locations in Japan and a few overseas sites is planned to give researchers access to central lab facilities, such as a massively parallel computer and a shared information database, as well as to provide an electronic notice board, electronic mail, and electronic meeting capabilities.

Theoretical foundation

The Feasibility Committee articulates the theoretical foundation of the RWC effort as follows:

> The real-world environment (including humans within it) is full of uncertainty and changeability. Human beings, however, can recognize the environment, make decisions, and act quite flexibly by processing in an integrated way a variety of information that is usually incomplete and ambiguous.
>
> Such information processing by human beings can be characterized by the terms *flexible information processing* or *real-world computing*, in contrast to the conventional hard information processing by computers, which assumes completely given information in a preassumed world or problem domain.
>
> The objective of research and development in the theoretical part of this program is to lay the theoretical foundation for and to pursue the technological realization of humanlike flexible information processing as a new paradigm of information processing.
>
> So far, much theoretical research has been done in fields related to flexible and parallel distributed information processing, for example, in pattern recognition, multivariate data analysis, probabilistic and statistical inference, fuzzy logic, neurocomputing, machine learning, regularization, and various optimization methods.
>
> To provide a theoretical foundation for flexible information processing, it is important not only to continue in-depth study in these research areas but also to clarify the theoretical framework of "soft logic" (flexible logical inference) commonly underlying these fields and to construct a new unified theoretical base for dynamic real-world computing. There, probabilistic and statistical formulation of problems and nonlinear dynamics in conjunction with learning and self-organization will be a key approach.

lem for which they have no immediate solution. European Community countries are still hesitant about participating.

In what follows, I closely paraphrase parts of the draft master plan. I focus on the research and development parts because the draft gives excellent (if overly long) lists of research topics.

Goals

The draft identifies two fundamental goals: integration of multimodal information (and of heterarchical processing modules), and learning and self-organization (optimization and adaptation). These system characteristics must be implemented in a generalized and flexible framework of massively parallel and distributed information processing. Researchers must also consider deeply the characteristics of real-world computing systems: openness, robustness, and real-time operation.

Research topics

Certain theoretical research topics are important for all the application domains and implementation methods.

Flexible representation of information. One goal is to treat in a unified manner information ranging from raw sensory data such as visual images and auditory signals to high-level symbolic languages. This requires a flexible framework for representing various kinds of information efficiently and effectively. The framework should be capable of representing certainty of information and suitable for implementing associative memory and learning or self-organization.

Important research topics in this area are:

- information representation based on multiway data tables;
- distributed and sparse representation of information;
- information representations suitable for implementing specific processing procedures;
- representation of probabilistic knowledge, stochastic automata, hidden Markov models, random Markov fields, probabilistic decision lists, and so on; and
- representation of information and knowledge as constraints.

Evaluation of information and processing models. Real-world computing systems must interact actively with the real world and learn or self-organize from experiences. To do this, they need a systematic framework for evaluating the importance of information and the capability to process models. Real-world systems also need evaluation criteria for regularization conditions used in information integration.

Research is needed to develop ways for systems to evaluate

- input and output information for learning or self-organization,
- regularization conditions for information integration,
- processing models and agents for cooperative integration of models, and
- input information for active interaction with real-world environments.

Researchers also must determine how to evaluate the flexibility and reliability of systems.

Flexible storage and recall of information. The brain's highly sophisticated memory functions permit flexible information processing. Real-world computing systems should have flexible memory functions to store and associatively recall

various kinds of information. Thus, researchers should analyze associative memory and develop new efficient mechanisms for flexible associative memory.

Research is needed in

- association using probabilistic reasoning,
- association using structural similarity,
- associative memory for storing time series, and
- associative memory using nonlinear dynamic systems.

Integration of information and processing modules. The draft views inference, prediction, and planning as the integration of various kinds of information and knowledge. It proposes theoretical research to analyze these information integration processes and develop a flexible way to control them.

Multivariate data analysis is a candidate approach for information integration. However, most conventional methods are limited to linear transformations; nonlinear extension will be important. Neural-network models give a kind of nonlinear extension. Regularization theory will provide a theoretical foundation for incorporating various kinds of constraints.

Cooperative processing by a huge number of processing modules is also important. The draft proposes research on the integration of processing modules and the following topics:

- integration of information using multivariate data analysis methods and neural-network models;
- integration of various kinds of constraints under the regularization theory;
- constraint logic systems, including dynamics for controlling information integration; and
- integration of information in human cognitive processes.

Learning and self-organization. According to the draft, the theory of learning and self-organization is crucial to constructing complex heterarchical systems for real-world computing. Memory and databases should also be self-organizing.

Important research topics are

- algorithms for learning probabilistic knowledge and heterarchically structured knowledge, for learning in a changeable environment, and for learning from uncertain information,
- learning algorithms using active information acquisition,
- methods for incorporating existing knowledge with the learning process, and
- selection of models for learning.

Optimization methods. The draft views information integration processes, learning, and self-organization as optimization processes. Solving such optimization problems requires a huge amount of computation. To lessen the computation required, researchers should develop approximately correct optimization methods for massively parallel systems.

Important research topics here are

- probabilistic optimization methods such as simulated annealing and genetic, ecological, and evolutionary algorithms;
- optimization using nonlinear dynamic systems such as neural networks; and
- other nonlinear optimization techniques.

Memories and databases should be self-organizing.

Novel functions for applications

Real-world computing systems support human activities by acquiring such information as images, speech sounds, and tactile input, and processing it for understanding or planning of controlled actions. The information is massive and modal. Moreover, real-world information is incomplete and uncertain.

Real-world computing requires novel functions with flexibility: robustness, openness, and real-time operation. The draft anticipates a broad range of application fields: recognition and understanding systems, flexible information bases (databases and knowledge bases), decision-supporting systems, flexible problem-solving systems, friendly man-machine interfaces, adaptive and large-scale simulation systems, and autonomous and cooperative control systems.

Novel-functions research should find an efficient way to integrate symbol and pattern, giving new kinds of robust functions. Merely combining conventional technologies or making ad hoc systems for specified tasks is not sufficient.

The draft identifies two kinds of integrated systems:

- real-world adaptable autonomous systems that understand and control the environment through active interaction with the real world, and
- information-integrating interactive systems that support and enhance human capabilities through enlarged information channels between humans and systems.

Adaptable autonomous systems. The draft calls for systems that operate between computing modules and the real world, partially replacing human activities in the real world. These systems will understand the real world and learn the concepts from real-world information necessary for adapting to new situations. Moreover, they will act to control the real world to obtain information and create desired situations. These

systems require novel functions to cope with the uncertain, incomplete, and variable characteristics of the real world.

Realizing real-world adaptable autonomous systems requires the following novel functions:

- capability to understand scene images,
- spoken natural language for questions and answers between the autonomous information system and humans,
- capability to plan for action sequences and adapt to the environment, and
- integration of pattern recognition, incremental world models, and control mechanisms.

Computers should understand conversational speaker-independent speech.

Information-integrating interactive systems. The second type of system operates between computing modules and humans. Humans handle various kinds of information and create information in which their intentions are embedded ambiguously. Computing modules are required to understand and integrate such information flexibly to assist humans in solving problems and creating new information. The systems should be regarded as powerful tools to extend humans' intelligent activities.

Realizing such systems requires these novel functions:

- capability to understand intentions from various kinds of information produced by humans,
- intelligent and interactive assistance to retrieve and present valuable information from a large amount of data in databases,
- intelligent simulation to create new information findings and forecast in the transient state of the real world, and
- integration methods for combining human factors and the real-world computational model.

Research topics. The computational bases for realizing the novel functions required by these systems will be massively parallel and distributed systems including neural systems or optical computing systems. The draft elaborates on the research necessary to realize both systems.

Flexible recognition and understanding. New paradigms for image understanding are needed for image processing with insufficient and inconsistent information. The early stages of image recognition and understanding (colors, shapes, position, and movements) are defined clearly in modules that can be formulated as optimization problems. These could be integrated, for example, according to the constraint-satisfaction paradigm. Systems need active conception to acquire the structure of image data. Sometimes introducing constraints in images improves the structure. Image processing then includes the computation of constraint satisfaction.

Flexible image understanding will emerge from novel algorithms for the segmentation of moving images and the integration of segmented objects by assuming the world model.

Important research themes in the flexible understanding of images include

- algorithms for learning about, self-organizing, and segmenting image data;
- integration of image features and conceptual symbols for understanding images;
- construction technology for handling a dynamically changing world based on image understanding;
- recognition of moving images of humans for understanding human intentions based on a flexible and deformable model for nonrigid object representation; and
- basic mechanisms for extracting intuitive information from images.

Speech is an important information source requiring flexible processing. Computers should understand conversational speaker-independent speech. Novel algorithms are needed to extract dynamic speech features and integrate recognition modules, including matching methods under the constraints of incomplete recognition.

Research is needed in the following aspects of speech understanding:

- early auditory processing and extraction of dynamic features of speech,
- mechanisms that can manage unknown words for the system, and
- simple and clear recognition algorithms suitable for massively parallel computation.

Technologies to make interfaces between computers and users based on flexible spoken and synthesized speech in the uncontrolled real world require research of

- robust processing of syntax and semantics for symbol sequences with noise;
- conversation models driven by speech and dynamic models of the world linked to conversational speech; and
- relationships between speech and other information sources (such as vision), which have complementary characteristics.

Very little of the large amount of raw and coded data in natural language is transformed by machine processing. Novel functions in this field should realize robust understanding systems that can manage a large amount of coded natural language as raw data and extract concept expressions for general purposes. Research in the following is needed:

- robust parsers applicable to a large amount of incomplete sentences,
- methods for embedding electronic dictionaries of different kinds into natural language understanding systems,
- self-organization algorithms for a sequence of sentences obtained from dialogue,
- computational models to integrate knowledge units through interaction based on conceptual and structural distance,
- explanation-oriented understanding of natural language, and
- interaction and complementary relationships between natural language and other types of information.

Flexible inferencing and problem solving. A flexible inferencing capability deals with incoherent or incomplete data created in the real world. Most data are expressed as symbols. However, inference systems may use numerical expressions such as probability and likelihood values. In this field, knowledge representation becomes important because it can formulate problem statements for users based on incomplete requirements. Flexibility of inferencing will require cooperation among inference agents; each agent will carry out a specified task.

Many large systems in our society, including production systems, have many different agents whose cooperation makes the global system stable, semioptimal, and adaptive. Problem solving is also important in a large system. Research to represent these mechanisms in a way appropriate to a massively parallel and semiuniform architecture is required to apply novel functions to the real world. This area requires research in

- semiglobal relationships or constraints and problem descriptions based on incomplete requirements,
- flexible problem-solving algorithms based on analogous or stochastic inference,
- cooperation among inference engine agents or with humans for realizing flexible inferences, and
- problem solving in large systems under the constraint of limited communication channels.

In their daily lives, humans are surrounded by different kinds of information: images, speech, text, tactile sensations, and so on. Some are produced and controlled; others are not. Humans use available input data for information retrieval by integrating different information and using flexible inference functions unconsciously. This mechanism must be determined to implement solutions in computer networks, since different kinds of data are distributed at different locations in our society. Integrating these data makes them more valuable. The flexibility of information bases is strongly connected with their self-organizing mechanism.

Research in the following is important for flexible information bases in

- knowledge representations common to or independent from each other,
- self-organizing mechanisms for large-scale information bases,
- learning algorithms for cooperative processing of many kinds of information and the acquisition of correspondences,
- formalization of a suitable interaction between each kind of information and the data structure, and
- inference for information retrieval.

Flexible human interface and simulation. A proposed novel information processing environment should provide users with a wide range of intelligent and physical collaborative activities with the computer. The user would interact with the computer through gestures, facial expressions, and spoken language, and receives information through real-time three-dimensional images. Realizing such an information processing environment requires research in

- broadband multimodal interfaces (using body action, facial expressions, and linguistic and visual patterns),
- information display using virtual reality, and
- cognitive and behavior models enabling systems to understand human intention by gesture and facial expression.

Real-world computing requires novel simulation technologies to provide a user with powerful tools to solve very difficult problems. Instead of doing expensive and time-consuming physical experiments that may not be precise, the user can turn to a computing system that simulates very large scale complex systems and predicts their behavior in real time. Prediction of untapped phenomena and future events—for example, in weather forecasting—is an important application of real-world computing. Research themes here are learning and adaptation simulation, and prediction and control of complex and chaotic time series.

Flexible autonomous control. Research is needed to realize a flexible autonomous coordinated system operating in the real world in real time—for example, a robot. Another application is aids for elderly or physically handicapped people. Such a system would have interacting function modules for perception, decision, and action. Problems to be solved

are how to integrate the functions and how to control interactions among function modules to achieve desired goals in the changing and ambiguous real world. Research here would consider

- flexible modeling of the environment, the task, and control;
- active, distributed sensing and sensory integration;
- on-site planning and distributed cooperative search;
- structure and coordination of multiple sensing, planning, and action modules for real-time skillful manipulation of objects; and
- maintenance of consistency between the internal world model with the dynamic real world.

Massively parallel systems

Real-world computing requires a computation framework that can process various kinds of information flexibly and with integrity. A system that implements applications with the novel functions just proposed is likely to consist of many modules and exploit parallel and distributed processing at several levels, both within and between modules.

Serious issues in massively parallel systems include the elimination of synchronization overhead, access contention, and communication overhead.

Several applicable parallel paradigms have been proposed, including concurrent object-oriented, dataflow, and data-parallel, neural-network, probability-based information processing. Real-world computing will probably be realized by some combination of these paradigms. Such a flexible information processing system will exploit many levels of parallelism. These new paradigms are naturally adapted to a massively parallel system, and they require a huge amount of computation to solve practical problems quickly. A massively parallel system that can efficiently execute multiple paradigms is necessary.

The massively parallel system should be flexible: in its hardware, its architecture, and its software. Hardware robustness and reliability are essential. A general-purpose architecture would support multiple paradigms. Software will be needed to realize adaptation, self-organization, and optimization.

Research topics. Achieving a general-purpose massively parallel system requires research in architecture, operating systems, languages, and system environments.

Architectures. Flexible execution models for general-purpose massively parallel architectures should fill the gap between the language models and hardware. Flexibility that allows a mapping of a virtual computer onto actual processing elements should be pursued. Architecture research must consider future device and packaging technology.

The interconnection network should provide high-speed communication comparable to the computation speed. It should also support dynamic load distribution, global synchronization, and global priority control. Not only silicon technologies but also optical technologies should be considered. Advanced intelligent routing, addressing, synchronization, deadlock prevention, flow control, and failure avoidance should be incorporated into a flexible network management system.

Hardware robustness to tolerate expected component failures in massively parallel systems should be examined. System components should have self-checking and self-repairing facilities. The overall system should have a maintenance architecture or facilities to maintain system reliability. Resource management should handle component failure with advanced intelligent routing and multiple-route processing.

The system's general-purpose nature may be achieved by mapping the multiparadigm model to the hardware execution model in software. If the hardware can be modified and adapted to a variety of applications, it may be considered as general purpose. Because a huge number of modules or processes must operate concurrently in real-world computing, architectural mechanisms to support rapid process switching and cheap synchronization are necessary.

Operating systems. An operating system for the massively parallel computer should support concurrent execution of various processes with high throughput. Simultaneous execution of parallel programs requires the operating system to partition resources among the programs and dynamically repartition at runtime. Locality, the concept of grouping activities, and the principle of balancing are guidelines for partitioning.

Realizing a functionally distributed system for flexible processor management may require the operating system to have a hierarchical structure much like human society. Hierarchical structure makes the system scalable for a massively parallel system. It requires efficient mechanisms to control activities and maintain the hierarchy in hardware and software.

In the massively parallel system, the elimination of synchronization overhead, access contention, and communication overhead will become serious issues. The operating system should be able to collect management information autonomously and undertake statistical or adaptive management. Memory management and virtual systems for several resources should be considered for efficient scheduling and

load distribution. For example, paging and virtual memory are the most successful approaches for memory management in conventional computers. It seems reasonable to expect benefits from these approaches in parallel systems. To balance the load among processors, designers may find it desirable to migrate a group of processes to another set of processors. Efficient hardware support and an operating system facility for migration are needed.

One function of an operating system is to provide programmers a higher level of abstraction to make parallel programming easier. Through the abstraction, the operating system manages the resources and coordinates several user programs.

The development of neural models is an important goal of the program.

Languages. The language for the massively parallel system must describe the coordinated operations of a number of processes. The draft proposes languages based on several paradigms. The problem is to extract the available parallelism in the problem domain and execute the problem with as much parallelism as the underlying system can provide. Development, verification, testing, debugging, and maintenance of large-scale parallel user programs will be complex compared with their sequential counterparts. Compilation and runtime implementation techniques scalable to nearly one million processors should be studied. The language model must be designed and prototyped, and the high-level language mode created.

The language model must describe flexible programming languages for massively parallel systems. It should provide a powerful architecture abstraction for software programmability, portability, and reusability, but be close enough to the architecture to retain computing power. According to the draft, a viable candidate is the object-oriented concurrency model. Inheritance adapts the objects to a certain computing environment. Reflection in this model makes the representation of the underlying program execution scheme manipulable within the language.

The primary features of high-level languages for the massively parallel system should be ease of programming and the capability to describe computation on the scale of a million processors. A viable candidate is an appropriate amalgamation, or layering, of concurrent object-oriented, functional, and declarative constraint-based approaches. Efficient interlayer compilation techniques must be developed. Speculative computation within the parallel constraint-satisfaction framework should also be exploited.

Because concurrent object-oriented languages are not intended to handle more than a million processes, extensions will be needed:

- a description system permitting hierarchical decomposition of complexity,
- diversification of message-propagation systems,
- self-reflection functions for adapting and evolving objects, and
- declarative descriptions of object relationships.

Environment for system development and programming. To give users a choice among various programming languages, the draft calls for a programming environment that can support multiple-paradigm programming.

System evaluation. Application fields will be picked to verify the effectiveness of massively parallel computation for real-world computing. An example is a simulation program to predict future events from established models. Such programs can predict macrobehavior by describing interactions between microelements. RWC organizers expect massively parallel systems to handle nonlinear or many-body problems by a direct mapping paradigm. They also expect the systems to handle simulations where the governing equations are difficult to formulate because of the complexity of the phenomena.

Neural systems

The brain performs certain computations many times faster than the fastest digital computer, even though the processing speed of neurons is five or six orders of magnitude slower than that of silicon logic gates. Neural systems are deliberately constructed to use some of the organizational principles in living creatures. The RWC draft proposes a brainlike neural system to perform special types of information processing at high speed and generate intelligent behavior.

In neural-network information processing, many simple processing units function according to the principles of "cooperation and competition" while maintaining close information exchange. In addition, since the programs used in conventional computers are expressed by the connection strength, a neural system can learn and adapt using conventional numerical optimization. These features make neural systems promising platforms for flexible information processing.

Recently, various neural-network models have been demonstrated as suitable for pattern processing, for finding approximate solutions to optimization problems, and for constraint-satisfaction processing. However, these networks have been restricted to small-scale applications. The neural

Overview of the RWC program schedule

According to the Real World Computing program draft master plan, the final massively parallel system will have on the order of a million processing elements. Because developing such a large system in one step is risky, the draft calls for prototyping. In the first half of the program, a prototype massively parallel system with 10^4 processing elements should be developed. However, several hardware prototype systems should be implemented, with one as a platform for software development and research on novel functions. Fundamental research on massively parallel models and architecture will be pursued concurrently. Various compilation and runtime implementation techniques will also be studied on the prototype systems, as well as a new user programming environment for millions of concurrent objects executing simultaneously.

After the first half of the program, these prototype systems and the platform system will be evaluated. The final system (or systems) designed in the second half of the program will be based on the prototype evaluation. The implementation will also be based on the research into theoretical foundations and novel functions in the first half of the program. A massively parallel system with 10^6 processing elements will be developed to execute various kinds of real-world computing applications in real time.

Many researchers in the program should be able to use the platform system without leaving their research institutions. A high-speed global network will connect the institutions and the platform system. In addition, the system software will support a standard user interface and a currently used system environment. RWC sponsors see facilities allowing many researchers to work in common circumstances as important in the promotion of the program.

models were very simple, and the learning method was almost exclusively back propagation.

The implementation of flexible information processing is difficult compared with past applications. The draft calls for the exploration of new capabilities in neural networks and new neural models. Today, neural networks are usually built and simulated on conventional computers, in which simulation speed is very slow, particularly with large networks. The problem may be overcome by using general-purpose massively parallel computers, but considering the cost and integration scale, special-purpose hardware will probably be needed.

The development of neural models is an important goal of the RWC program. The models must support large-scale neural networks that can acquire knowledge or adaptively change their structures by interacting with external environments. The hardware system should support a one-million-unit neural network and attain 10 tera connection updates per second (TCUPS). The final neural system will consist of special neural hardware and its software, and it will cooperate with the massively parallel system for flexible information processing. Integration will require cooperation with the research groups working on the massively parallel systems.

Research topics. The neural network will grow dynamically by supervised learning or self-organization in a real-time environment. However, the network should not forget existing knowledge as it grows. Thus, the network should be structured to have modularity, hierarchical structuring, and differentiation characteristics. Realizing these characteristics will require research in several fields.

Neuron unit models. To date, neural-network research has mainly used simple neuron unit models and has succeeded only in limited areas of application. The development of more sophisticated neuron models will lead to the generation of functional intelligence. Chaotic neuron models, complex-number neurons, and neuron logic models (in which neurons possess logical properties) have been proposed as candidate models. Their capabilities and the development of new neuron unit models should be evaluated.

Modularity and hierarchical structure. To achieve a system that acquires and restructures knowledge without destroying existing knowledge, researchers should consider the following:

- differentiation and hierarchical structure by adapting to the external environment,
- autonomous growth of neural networks by interaction with the real world,
- modularization using centered or distributed control,
- interaction between modules, and
- evaluation methods in modularization and hierarchical structuring.

Learning and self-organization. Many large-scale neural networks contain recurrent loops. Recurrent neural networks are promising for temporal pattern recognition and generation, while layered neural networks are effective for spatial pattern recognition. Also, recurrent neural networks can solve optimization problems because of their constraint-satisfaction capabilities. Recurrent neural networks should play an important role over the next 10 years, so learning and self-organization mechanisms for recurrent networks must be established.

Learning procedures are needed to construct optimal self-organized networks. Such networks should be sufficiently

large to learn the problem and small enough to generalize well.

Associative memories. Memory implemented by neural networks has interesting characteristics. Information such as spatial or temporal patterns can be stored using a distributed representation. This information can be retrieved using the closest match mechanism. Memory characteristics such as capacity and topological structure are important research topics for the clarification of the association principles and the development of engineering mechanisms to implement them.

Analog computation models. Information processing of neural systems can be supported by their analog nonlinear dynamics. Research should clarify new principles of analog computing in neural systems, including chaotic dynamics.

Integration of models. Models to integrate different paradigms should be investigated. The neural-network and logical processing could be integrated, for example, using a soft logic inference model. A combination of pattern and symbolic processing may be required to implement the neural system. The integrated model should run in a parallel processing environment, so theoretical and experimental research on input and output representations to and from neural networks is important.

Hardware. The one-million-unit neural network that interacts with the external environment and changes its structure adaptively may be modularized and consist of subneural networks, each of which has a thousand fully interconnected neurons. Achieving such a system requires that the final hardware must attain a 10-TCUPS processing speed. The hardware must process a wide variety of neural-network structures, such as the layered and recurrent types. General-purpose and scalable mechanisms should be embedded in the architecture.

The draft classifies neural hardware into three types:

- *Neuroaccelerators.* These special-purpose parallel processors are tailored for neural-network processing. Although many architectures have been proposed, the typical structure consists of hundreds of processing elements and achieves 1 GCUPS.
- *VLSI neurochips.* Hardware neuron units simulate neuron unit models. The spectrum of neurochip architectures is wide, ranging from digital processing chips to analog processing chips.
- *Engineering implementation of neural networks.* Hardware logic simulates the functions realized by neural networks.

According to the draft, it is impossible to compare these approaches, because each has its own features. The RWC program effort will focus on the second and third approaches.

Digital-circuit neurochips offer substantial advantages in noise tolerance and high-accuracy processing, and are suitable for the stable operation of large-scale systems. The manufacturing technologies for computers can be directly applied to neural digital circuits. In addition to the conventional approach, variations such as the pulse-density model are possible. New approaches may give the neurochip additional capabilities. Analog-circuit neurochips reduce hardware volume because operational circuits are fewer—an attractive feature in large-scale networks. Also, analog circuits can process dynamic and other complex neural networks such as chaotic neurons. Digital-analog hybrid chips may provide the advantages of both types of circuits.

Neurochips implement the neuron unit model directly. However, neural-network functions can be realized using logical hardware that does not simulate the neuron unit models. Such systems can be constructed by the repetition of simple logical circuits.

Achieving a 1-million-unit neural network means the hardware must process at 10 TCUPS.

In a neural system, all the units exchange their activation values. RWC program research will consider time-multiplexing and frequency-multiplexing methods for the interconnection network. Interconnection network design is closely related to the implementation technology. Bandwidth is restricted by neurochip and board pin number. Wafer-scale integration, three-dimensional VLSI architecture, and optical interconnection will be important. For neural application-specific ICs, tools such as libraries of standard neuron cells will be required. Silicon compilation of networks is important for embedded applications.

Software. Most neural-network researchers have their own simulators. The draft calls for a flexible general-purpose neural simulator for large-scale neural networks. The requirements for such a simulator are machine independence, expandability, a convenient user interface, high-speed processing, and a variety of utility routines. Analytic tools should be integrated into the simulator to describe the algebra and geometry of neural learning and to permit transformations of the network dynamics to achieve better convergence and cognitive performance.

Neural-network processing should be programmed with high-level languages, so research is needed in the

- representation of ambiguous information,
- description of best match operations,
- integration with logical programming, and

- integration with simulators.

In general, the number of hardware neurons will be smaller than the number of units in a neural network. Therefore, a virtual mechanism will be needed to fill the gap.

Integration with massively parallel systems. The neural system may consist of various neural subsystems. However, the system will be a special-purpose processor for pattern recognition, associative memory, optimal combinatorial search, and other special functions. Because it will be connected to or integrated with other systems, integration mechanisms to be researched include

- *Closely coupled integration.* Neural networks are connected to each processing element of a massively parallel system as an associative memory.
- *Loosely coupled integration.* A massively parallel system sends a special job such as an optimization problem to the neural system as a dedicated system.

System evaluation. According to the draft, neural theories and models will be evaluated using two methods for the measurement of processing speed (TCUPS), and evaluation of the execution of real applications. The benchmark application should require flexible and high-speed information processing. Candidate applications are motion image processing and information integration by multineural networks.

THE DRAFT CONCLUDES HERE WITH VARIOUS COMMENTS on plans for optical computing systems. Because of space limitations, I do not include them here.

Questions regarding this column can be addressed via e-mail to David K. Kahaner, US Office of Naval Research, Far East, at kahaner@cs.titech.ac.jp.

Reader Interest Survey

Indicate your interest in this article by circling the appropriate numbers on the Reader Service Card.

Low 171 Medium 172 High 173

IEEE MICRO

Editorial Calendar

OCTOBER 1992

Processing hardware for video communication

- ICs for HDTV compression
- ICs for HDTV communication
- Parallel processing systems with real-time video compression capabilities
- Multimedia systems with real-time video compression capabilities

Ad closing date: September 1

APRIL 1993

Advanced packing and interconnect techuology

- Critical packaging trends and issues
- Substrate and package technologies—for example, flexible, glass, or diamond substrates; few-chip or 3D packaging
- Attachment, bonding, and connection technologies, including five-pitch surface mount, laser applications, known-good die, and interconnection trade-off analysis

Ad closing date: March 1

DECEMBER 1992

Special signal processors

- Recent information from the vital area of digital signal processors
- News about other techniques for signal processing
- Mixed analog/digital processors solve a real need
- Neural networks process signals following methods used by the human brain

Ad closing date: November 1

JUNE 1993

Hot Chips IV

- This extremely popular issue presents the latest developments in microprocessor and chip technology used to construct high-performance workstations and systems as presented at the annual IEEE Computer Society TCMM-sponsored symposium

Ad closing date: May 1

FEBRUARY 1993

Automotive/traffic microelectrouics

- Worldwide developments in microelectronics for traffic and driving assistance
- Improving traffic safety with electronics
- Latest developments from Japan, the European Prometheus, and US IVHS programs

Ad closing date: January 2

AUGUST 1993

General Interest

- Multiprocessing
- Optical and biological computing
- Microcomputing to aid the handicapped
- Systems design
- DSP-based tools

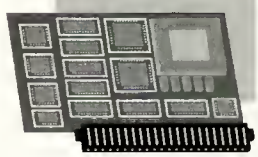
Ad closing date: July 1

IEEE Micro helps designers and users of microprocessor and microcomputer systems explore the latest technologies to achieve business and research objectives. Feature articles in *IEEE Micro* reflect original works relating to the design, performance, or application of microprocessors and microcomputers.
All manuscripts are subject to a peer-review process consistent with professional-level technical publications. *IEEE Micro* is a bimonthly publication of the IEEE Computer Society.

Advertising information: Contact Marian Tibayan, Advertising Department, IEEE Computer Society, PO Box 3014, Los Alamitos, CA 90720-1264; (714) 821-8380; fax (714) 821-4010.

Articles may change. Please contact editor to confirm.

On the Edge



Electrical engineer registration

[This issue, Michael Lindeburg advocates registration for electrical engineers. I invite readers to send information on a tool or method that solves problems, for consideration in future issues. – C.W.]

Michael R. Lindeburg, PE
Professional Engineering Institute

Carl Warren

McDonnell Douglas

Space Systems Company

(714) 896-3311

x. 7-1230

warren@ssdunx.mdc.com

Although the number of electrical engineers pursuing professional registration in the US is on the rise—4,225 took the Professional Engineer (PE) examination in 1991, compared to 2,698 in 1987—the conventional wisdom among most EEs is that registration is unnecessary. Though it is difficult to pinpoint how many EEs are registered in the US, a recent IEEE survey listed only 20 percent of its approximately 325,000 members as PEs.

So why haven't the other 80 percent registered? To many EEs, professional registration is an aspect of the profession they're vaguely aware of, but one that doesn't really apply to them. It's as if registration is something other people do.

One reason engineers feel this way is because most states waive required registration for engineers working in industry. This exemption contributes to the low registration numbers, since so many EEs are employed by the private sector. Generally, states require registration only for engineers who consult, work on projects concerning public safety, or testify as expert witnesses in court—a profitable endeavor.

For engineers who do not fall into one of these categories, registration may seem unnecessary. Thomas Holden is a Silicon Valley telecommunications specialist with a master's degree in electrical engineering from the University of California at Davis. He says registration would make little difference at this point in his career, except in professional prestige. Although registration isn't in his immediate plans, Holden plans to pursue his professional license later so he can sign off on independent consulting projects, one of the greatest benefits of registration.

In certain industries registration usually results in greater career advancement and higher salaries. Holden says he has consistently seen engineers who work for large defense contractors and utility services move more quickly up the career ladder if they are registered.

While some companies offer career advancement incentives to registered engineers, other employers require registration due to the number of foreign-born engineers working in the US. Without the PE license, it is sometimes difficult for companies to assess the quality of education the engineers received in their native countries.

Promoting the profession

In addition to corporate requirements for a professional license and the financial advantages enjoyed by registered engineers, people often overlook a greater overall benefit: Registration enhances not only the electrical engineering community but the entire engineering profession.

Professional registration is the first step in promoting the profession. Many engineers believe mandatory registration would encourage greater responsibility, enhance their public reputation, and lead to greater financial rewards for engineers in all disciplines.

Oliver C. Morse, a registered electrical engineer and PE review course instructor in Berkeley, California, believes engineers will not achieve their rightful professional and financial status until they take registration more seriously.

0272-1732/92/0800-0082$03.00 © 1992 IEEE

Morse, a plant engineer specializing in energy-efficient lighting at Lawrence Berkeley Laboratories, says making registration for engineers a universal requirement—especially in electronics—would be a major step in enhancing the reputation and rewards of the profession.

"If engineering registration isn't considered important and the attitude toward it remains lackadaisical, the status quo won't change," says Morse. "The continued lack of emphasis on registration will preclude improvement."

Morse believes someone needs to step in to push for required registration for engineers. He sees the National Society of Professional Engineers as the group that could do it.

If the NSPE uses its considerable influence to unify the individual societies of each engineering discipline, he says, mandatory registration could one day become a reality that would "upgrade the profession."

History

Upgrading the engineering profession was one of the original intents of instituting registration. To better understand registration, let's look at what registration actually is and why it came about.

Engineering registration in the US is the examination process by which a state's board of engineering licensing (the registration board) determines and certifies a minimum level of competence. This process protects the public by preventing unqualified individuals from offering engineering services.

Professional registration for engineers was also motivated by concern for the public welfare. In the US, registration began in turn-of-the-century Wyoming, where almost everyone (except, it seemed, engineers and land surveyors) prepared maps and plans of homesteads, streams, canals, and water works. These maps and plans were often grossly inaccurate, causing property losses and other problems.

Wyoming's State Engineer, Clarence T. Johnston, corrected the situation by initiating legislation to require all engineers and land surveyors to register with the state. His bill passed in 1907.

As awareness of the public responsibility of engineers for designing buildings, roads, and bridges began to grow, other states enacted similar registration requirements. Today, all 50 states (plus the District of Columbia, Puerto Rico, Guam, and the Virgin Islands) regulate the registration and activities of engineers.

However, due to industrial and other exemptions, most engineers do not need to register; fewer than one third of the engineers in the US have. Statistics from the National Council of Examiners for Engineering and Surveying (NCEES) list about 340,000 engineers nationwide registered as of 1990 in at least one state.

No one has tried to create a national system of registration, so requirements and procedures vary among the states. However, because the states recognize the need for uniformity, most use similar tests.

The uniformity results from NCEES's efforts. NCEES produces, distributes, and scores the national Fundamentals of Engineering (FE) and PE exams. States purchase the exams from NCEES and administer them. NCEES does not distribute applications to take the exam, administer the exams or appeals, or notify examinees of their results. The individual state boards of engineering registration perform these tasks.

Exams

Most states share similar testing procedures: Applicants take two eight-hour

Resources for US registration

Organizations

National Council of Examiners for Engineering and Surveying (NCEES)
PO Box 1686, Clemson, SC 29633; (803) 654-6824

National Society of Professional Engineers (NSPE)
1420 King St., Alexandria, VA 22314; (703) 684-2800

Professional Engineering Institute
1250 Fifth Ave., Belmont, CA 94002; (415) 593-9731
A nonprofit educational organization that provides assistance, study materials, and state-approved review courses to all engineers seeking registration.

Professional Publications, Inc.
Address same as above; (415) 593-9119
Specializes in review and reference books for engineers, architects, land surveyors, and interior designers preparing for professional licensing exams.

Books

M.R. Lindeburg, *Engineer-in-Training Reference Manual*, 8th ed., Professional Publications, Belmont, Calif., 1992; $45.95

D.G. Sunar, *How to Become a Professional Engineer*, 3rd ed., Professional Publications, 1991; $10.95

R.B. Yarbrough, *Electrical Engineering Reference Manual*, 5th ed., Professional Publications, 1990; $45.95

written exams. The first is the FE exam, also called the Engineer-In-Training (EIT) exam. This exam covers basic subjects from mathematics, physics, chemistry, and engineering classes from the first years of college.

To qualify for the FE exam, applicants must meet the most standard requirement, that is, graduation from a four-year program in engineering, basic sciences, or engineering technology accredited by the Accreditation Board for Engineering and Technology (ABET). In most states, students in their last year or two of an accredited program can take the FE exam. Graduates of nonaccredited programs may need to show acceptable work experience before they may take the exam.

After passing the FE exam and completing several years (usually at least four) of relevant qualifying experience, the next step is the PE exam, also called the Principles and Practice of Engineering exam. In most states, the examination covers only subjects from a specific engineering discipline. Table 1 lists subjects addressed on the electrical PE exam, and Figure 1 offers two sample questions.

Although the testing procedures among the states are similar, the required qualifications and the methods of application vary. Contact your state board for specific licensing procedures.

Table 1. Breakdown of problems on electrical PE exam.

Subject	Number of problems
Generation systems	2
Transmission and distribution systems	5
Rotating machines	1
Lightning protection and grounding	1
Control systems	2
Electronic devices	3
Instrumentation	3
Digital systems	2
Computer systems	3
Communication systems	3
Biomedical systems	1

Situation. A shunt excited DC motor is rated 1,150 rpm, 10 hp, 80 amps, and 40°F temperature rise. The motor drives a load for which the torque is to be constant regardless of speed.
Requirements. Draw and identify the circuit elements for:
(a) a motor speed of 1,300 rpm
(b) a motor speed of 800 rpm

Situation. An electrical supply of 20 volts ±5 percent with 3 ohms internal resistance.
Requirements.
(a) Design a Zener diode regulator to supply a variable 15 to 130 ohm load with a constant 15 volts DC.
(b) Sketch the complete circuit and specify all components.

Figure 1. Sample problems. These problems are taken from *How To Become A Professional Engineer* and are not actual problems from NCEES exams.

The exam process may sound daunting, but passing rates are actually quite high. According to recent NCEES statistics, nearly two thirds of those who took the FE exam in recent years passed. The odds are considerably better for graduates of ABET-accredited engineering programs; passing rates run as high as 85 percent.

The PE exam is obviously more specialized and difficult than the FE exam, and requires more preparation and specific knowledge. Passing rates for EEs taking the exam for the first time ranged from 38 to 46 percent between 1987 and 1991.

State boards provide additional information on requirements and procedures of registration. Although no single reason may convince engineers to pursue professional registration, the overall benefits to the engineer and to the profession suggest that professional licensing can significantly enhance and broaden an engineering career.

Michael R. Lindeburg is founder and president of Professional Publications, Inc. and the Professional Engineering Institute, both in Belmont, California. Lindeburg received his BS and MS degrees in industrial engineering from Stanford University. He worked as a senior industrial engineer at GTE Lenkurt in San Carlos, California, before teaching his first FE review course at DeAnza College in Cupertino. He is a registered industrial engineer and a member of the NSPE and the American Society for Engineering Education.

Software Report

Melco's neural chip; holography research

David K. Kahaner

US Office of Naval

Research, Far East

kahaner@cs.titech.ac.jp

Western scientists, especially those in universities, often have only vague ideas about the real size of some Japanese high-technology companies. Take for example Mitsubishi, which is probably best known in the West for its automobiles and ships. I wonder how many realize that Mitsubishi Electric (Melco), which does not include either the automotive or ship construction companies, is itself a 71-year-old, $27-billion company with just under 100,000 employees.

Melco produces consumer products (audiovisual, home electronics, and appliances); information, telecommunication, and electronic devices and systems; industrial and automotive equipment; and heavy machinery, including locomotives and nuclear power equipment.

The consumer and information product divisions are nearly equal in value with about 60 percent of sales. Table 1 on the next page shows research and development expenses, mostly in electronics, new energy sources, new materials, and biotechnology. Melco has quite a few research labs, the oldest being the Central Research Lab near Osaka.

Melco is constructing a synchrotron-radiation facility. It has an active superconductivity program and is working on 64-Mbit DRAMs, a GaAs semiconductor laser and FET, an optical neurochip, a process using ultrafine ice particles as a cleaning agent in electronic device manufacture, semiconductor ceramic fiber (30-50 microns), and a very high integration digital signal processor. The company heavily supports advanced transportation, such as superconducting trains (magnetic levitation) and superconducting ship propulsion systems, and it supplies vast numbers of locomotives and rail stock to countries from China to Mexico.

Melco is active in the Institute for New Generation Computer Technology (ICOT) program with the development of several versions of a parallel sequential inference (PSI) machine, including the first one in Japan. One application is a Japanese-English translation system. Other parallel processing projects include a high-speed database machine and a parallel-syntax processor for a man-machine speech recognition system.

The company has produced more than 150,000 elevators and escalators. One group uses AI, expert systems, and fuzzy theory and has recently produced the world's first spiral escalator (installed in San Francisco). Similarly it has a fuzzy controller for electric discharge machines, an expert system for insulation diagnosis, and a developing knowledge media station.

Melco's value-added network (packet-switching) is Japan's largest with plans to link 20,000 terminals in its corporate e-mail network, including Unix workstations in technical divisions and PCs in administrative divisions. It will install a server in each lab's LAN and a main server in the main office. When completed this year, the network will connect 63 offices and 45 sales companies. The company hopes the network will reduce the paperwork that flows into its main office—currently three tons a day!

Melco is also developing several products that will support the Integrated Services Digital Network (ISDN) now being installed in Japan. For example, one home automation system monitors for gas leaks, fire, unlocked doors, and similar hazards. A user can start the system with a phone call.

Visiting the CRL. Wolfgang Banzhaf, a researcher visiting from the Institute of Theoretical Physics and Synergetics in Stuttgart, Germany, arranged my visit to Melco's Central Research

0272-1732/92/0800-0085$03.00 © 1992 IEEE

Table 1. Melco R&D costs (in billions of yen and dollars).

Year	Yen	$US	Sales (%)
1987	94.6	0.76	4.8
1988	101.9	0.82	4.9
1989	118.5	0.95	4.4
1990	145.1	1.16	4.3
1991	159.8	1.20	4.9

Laboratory. Banzhaf works in the neurocomputing group, run by Kazuo Kyuma. (Recently, Kyuma received the Sakurai prize, Japan's most renowned award for optical sciences, partially for his work on the optical neural chip.)

Kyuma's staff of 18 scientists have backgrounds in physics, applied mathematics, materials, and other disciplines. Five of them have PhDs, and three more may get that degree in the future. Typically, Japanese companies hire recent graduates with the equivalent of a bachelor's or master's degree; few recruit university-trained PhDs. Japanese scientists working at an industrial lab can get a PhD after several years and enough published research. Scientists in the US can also earn PhDs while working off campus. But they usually are well known to their faculty advisor and have a lot of interaction with the university. In Japan the student-university connection is much weaker.

I've often wondered how Japanese companies can produce so much interesting science with this kind of a system. One reason is that good companies are extremely selective and hire only a few students each year—with luck, the very best. These stay with their groups for a long time. Promising staff are often sent outside the company for training or additional education. Group and project leaders are more senior and usually have advanced degrees, often obtained via industrial publications. Sometimes industrial scientists leave their companies for academic positions. The good Japanese research labs are also sprinkled with Western visitors. (To be perfectly honest the equation still does not compute for me, but the results speak for themselves.)

The neurocomputing group can access Melco's Cray Y-MP in the Materials and Electronic Devices Lab, but the group does most of its simulation work on workstations, which they sometimes cluster. Given the large amount of simulation and the sophistication of the models, I am surprised they don't have access to a parallel machine, such as a CM-2. The Industrial Electronics and Systems Lab has an NCube, but the neural net group does not use it.

Neural chip. The most exciting project here is a fully optical neural chip with eight neurons and (optically) adjustable weights. Melco's chip is the most advanced to my knowledge. The company developed it using molecular-beam-epitaxy crystal growth and gallium arsenide opto-electronic device technologies. The first prototype, a one-foot cube, was developed in 1988. The current chip is mounted on a board together with other large-scale integration devices in a demonstration neural computer hooked to a small workstation. The chip has three stacked layers: an LED array, an interconnection matrix, and a photodiode array. The two arrays are placed in a crossbar with a dynamic spatial light modulator sandwiched between them as the interconnection weight matrix.

The original neural network used spatial light modulators. The new neurochip uses variable sensitivity of its photodiodes, controlled externally, to adjust the weights. This arrangement allows vector matrix multiplication to process in parallel, and because the weight matrix can be adjusted, it also means that the chip can learn. This is quite distinct from other approaches in which the interconnection weights are fixed in hardware and use Hopfield models without learning ability. The current chip is an 8×8 arrangement, and the group says its construction approach allows for about 2,000 neurons/cm^2.

A major ingredient of Melco's approach is quantized learning models. By that I mean that neural connection weights are not permitted to take on continuous values but are restricted to a finite set, often just two or three values. A quantized learning rule for back propagation is as follows:

1) Start with random weights (continuous).
2) Quantize into several discrete levels.
3) Set spatial light modulator to value of discrete (hardware) weights.
4) Present one of training and supervised signals to network.
5) Calculate error by conventional back-propagation method.
6) Correct continuous weight by adding error.
7) Repeat steps 2-6 for all training signals until connections converge.

In this approach, the quantized weights are built on the optical hardware and the continuous weights are stored in memory, allowing continuous changes in the weights and parallel optical implementation at the same time. Melco staff have shown that the chip can learn even if synaptic weights are quantized into only two levels. They have many simulation results and implementation experiments using binary operating SLMs to recognize 26 alphabetic characters.[1,2]

Group interaction. Several of the group members commented that they appreciated the opportunity to spend time talking to me and also listening over my shoulder while their colleagues described their own research work. In fact, they felt such interactions were very uncommon.

This surprised me since the group works in a large, airy room with modest partitions between desks, and I expected free interchange would be the rule rather than the exception. They told me they have plenty of joint work

but very little free time for researchers who do not work directly together on a project to discuss their work informally. An internal seminar was started last year, but hasn't succeeded too well. Work hours are for working, lunch is normally at one's desk, and seminars are held after 5 p.m. The scientists said they felt seminar times were symbolic and the atmosphere discouraged the kind of cross-group interaction typical at Western research labs or at universities.

The neural net group also develops algorithms and simulates different learning models. Some of this is clearly motivated by the parallel work on the optical neural chip just described.[3] This work tries to solve the performance degradation problem usually associated with weight quantization. (Performance here is the number of steps necessary for successful training.)

Stamp recognition. Another project seeks to identify the value of a postage stamp from its image.[4] The interesting wrinkle here is that the stamp image is not digitized. Rather, a small number (32) of sensors read intensity and frequency from parts of the stamp and their values are input to a neural net, which outputs the stamp's face value. The stamp recognition project has succeeded (at least partially) in optical implementation. The neural network has been implemented; only the sensor device has not (so far).

Symposium on 3D imaging

In February I attended the International Symposium on Three Dimensional Image Technology and Arts, sponsored by the University of Tokyo's Institute of Industrial Science, to see some of the most recent research in this field. Nearly 130 scientists participated (most from Japan, but some from North America and Europe), and 37 papers were presented, 17 from outside Japan.

Japan has an active 3D research community, and this symposium was supported by various Japanese electronics, television, and robotics societies, as well as the Ministry of Posts and Telecommunications. At the opening ceremony, an MPT representative explained that his ministry considers 3D their most important research project. A five-year project focuses on developing technology to transfer 3D images on optical fiber for TV telephones and conference systems.

Current work seems aimed at stereo vision, but the most promising technique for full 3D imaging is holography, the reconstruction of the object wavefront. Holography's original principle involves illuminating the object with a laser and simultaneously recording the reflected (or diffused) light from the object and a reference beam from the laser, creating an interference fringe pattern. The recorded pattern can later be illuminated with the same laser to reproduce the image.

Work in holographic techniques focuses on two areas:

- using conventional light rather than a laser, and
- creating holographic stereograms in which cameras at different positions record multiple images of an object while a hologram of each image is recorded sequentially.

MIT's media lab, under the direction of Stephen Benton, carries out some of the most exciting developments in this area. Benton's early claim to fame was the invention of the white light transmission (or rainbow) hologram and, more recently, holographic video. The lab is also working on large, full-color, animated, computer-synthesized holograms. Perhaps most importantly, this work has reenergized the field and forced researchers to look seriously at ongoing related work.

At this symposium, papers discussed holography in medicine, cameras for holographic imaging of moving objects, and holographic TV using LCDs. All the speakers pointed to Benton's work as years ahead of other work. Although he did not attend the symposium, his work is known from his publications and his recent visit to Japan.

Nevertheless we are still years away from holographic TV in our homes. Any practical holographic display device relying on Benton's approach will require a time-bandwidth product far exceeding those available with single channel acousto-optic modulators and require multichannel modulators and parallelism.

References

1. J. Ohta, Y. Nitta, and K. Kyuma, "Dynamic Optical Neurochip Using Variable-Sensitivity Photodiodes," *Optics Letters*, Vol. 16, No. 10, May 1991, pp. 744-746.
2. M. Takahashi et al., "A Quantized Back Propagation Learning Rule and Its Application to Optical Neural Networks," *Optical Computing and Processing*, Vol. 1, No. 2, 1991, pp. 175-182.
3. W. Balzer et al., "Weight Quantization in Boltzman Machines," *Neural Networks*, Vol. 4, 1991, pp. 405-409.
4. W. Banzhaf et al., "Optical Implementation of a Competitive Network," *Fast Intelligent Systems*, B. Soucek, ed., *Wiley Series in Sixth Generation Computing Technologies*, Wiley & Sons, New York, 1992.

Micro News

Send information for inclusions in Micro News one month before cover date to Managing Editor, IEEE Micro, *PO Box 3014, Los Alamitos, CA 90720-1264.*

Atomic fountains, laser tweezers, and optical molasses

Stanford University physicist Steven Chu says he and his coworkers are using lasers to hold and move atoms and molecules with exquisite control.

Chu has used laser beams to chill sodium atoms to within a tiny fraction of a degree of absolute zero and create "atomic fountains" of slowly falling atoms. According to Chu, these fountains allow such precise measurement that the scientists can detect the gravitational change caused by moving four inches farther from the center of the earth.

Chu also uses lasers as "optical tweezers" to grab and move individual biological molecules. Researchers can stretch out single molecules of DNA in water and observe them as they snap back elastically. They can also pin the stretched molecule onto a microscope slide for observation.

Though cooling atoms to temperatures near absolute zero by zapping them with laser beams seems contradictory, Chu uses a technique called "optical molasses" to do just that. The technique was proposed by Ted Hansch and Art Schawlow in 1975 and demonstrated by Chu and colleagues in 1985. Optical molasses makes high-tech use of the Doppler shift effect, which raises a laser beam's frequency slightly. When an atom is moving toward the beam, the atom absorbs a photon of light energy. The impulse delivered to the atom during its collision with the photon slows the atom's motion.

"Six years ago, none of these applications were considered," says Chu, who holds patents on some of the research along with Stanford University.

Lens focuses cold neutrons

A team of US and Russian scientists has developed a lens that focuses cold neutrons to better analyze elements in advanced materials. Semiconductor manufacturers will use this nondestructive method of measuring elements in materials from the wafer stage through assembled components.

In the past, scientists have used cold neutrons to measure the quantity and location of chemical elements in materials. A chamber of heavy-water ice at −235°C chills the neutrons to slow them from about 2,200 meters per second to a few hundred meters per second. Cold neutrons can indicate where dopants and contaminants lie in a silicon semiconductor. However, neutrons tend to diverge as they come out of a reactor and must be collimated with shielding onto the sample.

The lens represents a new way to control neutrons. It comprises 721 fibers, each containing a conical arrangement of 1,261 hollow glass capillaries that guide the neutrons into a convergent beam. A reactor produces the neutrons, which then pass through the cold source and lens to reach a focal point 104 mm beyond the lens. Researchers place a sample of the material in the beam's focus, where the neutrons are most concentrated to discern the most information about the material. With the more intensely focused neuron beam, scientists can get better quality information about elemental distributions and hope soon to produce 3D information without cutting into the samples.

Researchers at the I.V. Kurchatov Institute of Atomic Energy in Moscow and the Institute for Roentgen Optical Systems in Moscow designed the lens. Scientists at the US Department of Commerce's National Institute of Standards and Technology, working with X-ray Optical Systems at the State University of New York at Albany, used the lens to create the focused beam. According to NIST, manufacturers may in the future send samples of material to be tested using these lenses at one of about a dozen reactor centers in the US and use the resultant data for materials

0272-1732/92/0800-0088$03.00 © 1992 IEEE

development or to calibrate their internal measurement techniques.

IEEE approves SCI standard

The IEEE Computer Society's Standards Board has approved the Scalable Coherent Interface as IEEE Std 1596-1992. SCI is similar to a standard computer bus but operates at higher speeds by using an efficient packet-based protocol over a large number of independent, unidirectional point-to-point links.

Initial SCI links transmit data at 1 Gbyte/s (16 bits wide, differential) up to a few meters and 1 Gbyte/s (bit serial) over coaxial cable for 20 meters or fiber optics for several kilometers. Chips and transceivers that employ the SCI protocols are expected by the fourth quarter of this year.

(For more details on SCI, see "Scalable Coherent Interface and Related Standards Projects," *IEEE Micro*, April 1992, pp. 10-22.)

Prize finalists selected

Finalists have been selected in the 1992 Gordon Bell Prize competition for significant achievement in the application of parallel processing to practical scientific and engineering problems. "The high quality of the entries made the selection of finalists particularly difficult this year," said Alan Karp, prize administrator.

Two prizes of $1,000 each will be presented at Supercomputing 92, November 16-20, in Minneapolis. The two winners will be chosen from entries submitted in three categories: performance, price/performance, and compiler parallelization.

MOS transistor inventor

Dawon Kahng, inventor of the first operative silicon MOS transistor and founding president of the NEC Research Institute in Princeton, died recently in New Brunswick, New Jersey. The Korean-born physicist was 61.

During a long career at AT&T Bell Laboratories, Kahng also invented the floating-gate memory cell and contributed to the development of Schottky junction devices and electroluminescence. In 1984, Bell made him a Fellow in recognition of his outstanding research efforts.

After retirement, Kahng pursued long-term basic research in sciences underlying future computer and communications technology at NEC Research. He was a Fellow of the Institute of Electrical and Electronics Engineers.

Carl Warren leaves us

IEEE Micro Editorial Board member and columnist Carl Warren died July 18 after a long bout with cancer. Immediately upon joining the board in October 1989, Warren introduced the popular On the Edge column, revived Micro Standards, and helped review manuscripts. The McDonnell Douglas Space Systems senior scientist, member of several technical organizations, and former *Interface Age* editor-in-chief and *EDN* regional technology editor had made many friends around the world.

In recognition of his extensive support for *IEEE Micro*, the Computer Society awarded him the Certificate of Merit. Editor-in-Chief Dante Del Corso expressed his appreciation at the time of the award saying, "Carl Warren is one of the most reliable editors on our board. The two departments coordinated by Carl brought a lot of useful information to readers in these last few years."

Carl Warren will be sorely missed by all who knew him.

Micro bits

- IBM recently announced an agreement with Siemens of Germany and Toshiba Corp. to develop **0.25-micron 256-Mbit chips.** The three-continent alliance may spend $1 billion in joint production of the memory chips at IBM's US facilities.
- Analog Devices, Electronic Designs, and Westinghouse Electronic Systems Group signed a contract with MIT/Lincoln Laboratory to develop a **high-performance parallel DSP chip.** The CMOS device advances VLSI process technology with 0.6-mm geometry and 3.3V operation.
- Advanced Micro Devices joins with Fujitsu to build **flash memories** at a new factory in Japan. The two companies agreed to invest up to 5 percent in each other's stock.
- **IEEE Standards** offers full-text standards on CD-ROM with graphics to PC users. For more information, call 1-800-241-7824.

New Products

Send announcements of new microcomputer and microprocessor products to Managing Editor, IEEE Micro, PO Box 3014, Los Alamitos, CA 90720-1264.

Joe Hootman

University of North Dakota

Chips

Step down to 3.3V

Two surface-mount converters integrate 3.3V power into a system without noise or spikes, according to the manufacturer. The MP7101 one-amp, 5V to 3.3V DC/DC converter needs no external filter, regulates output, and limits output noise to 0.1 μV. The MP7008 incorporates these features into a regulator with thermal shutdown and current limit protection. It has a low-power standby mode, a wide input range, and an adjustable output version. Both devices accept inputs up to 35VDC and measure 1×1.5×0.5 inches. *Modupower; $17 each (100s).*

Reader Service No. 10

Low-voltage microprocessors

Five low-voltage derivatives of the 68HC11 family of 8-bit microcontrollers support power-sensitive applications, such as hand-held and portable systems. The chips (A8, E9, D3, L6, and K4) come in 3.0V to 5.5V versions. In addition, the manufacturer sells two lower pin-count versions, the KA4 and the 7KA4. The lower pin count allows the devices to be used in space-sensitive applications, such as cars, consumer electronics, and personal computers.

Features vary among versions. One vendor using the L6 has built a small, light-weight, digital, cordless telephone that runs for six hours on three AAA batteries. (See box on Motorola's other family of 8-bit microprocessors.) *Motorola; from $7.94 (10,000s).*

Reader Service No. 11

Performs at 1.2V

Two families of 3.3V ICs can extend battery life in portables and reduce battery cell requirements. Chips in the HLL (high-speed, low-voltage, low-power) family run twice as fast as bipolar 5V ICs. The submicron, CMOS-manufactured chips have a 2.5-ns propagation delay for an octal buffer line driver. The LV (low-voltage) family, with an 8-ns delay, supports similar applications that don't require as much speed.

Chips in both families perform with as little as 1.2V, the low end of NiCad rechargeable batteries before they need recharging. *Philips; price not given.*

Reader Service No. 12

Hard-drive chip set

A three-chip set of CMOS devices replaces up to eight components in hard-drive systems. The Reach2 read-channel device integrates the analog read channel, frequency synthesizer, and servo demodulator circuitry for multizone, constant density writing and reading. It supports data rates from 6.67 to 30 Mbits/s. Search1 contains the circuitry needed for the position and spindle-servo systems. It comprises three processors for mass storage: a general-purpose processor, DSP, and servo timing processor. Spin1, a DSP servo data converter, interfaces to an 8- or 16-bit microcontroller.

The 0.9-micron chips dissipate less than one watt in a 2.5-inch drive. The manufacturer says this power economy prevents overheating in drive subsystems smaller than 1.8 inches. *AT&T Microelectronics; $25 (Reach2), $20 (Search1), $10 (Spin1) (1,000s).*

Reader Service No. 13

Monitors digital circuits

The SN74ACT8994 digital bus monitor emulates logic and signature analysis test instruments. It lets designers using the JTAG (IEEE Std 1149.1) boundary scan architecture monitor digital circuits, without having to physically access the devices under test. The SN74ACT8994 synchronizes with the functional devices on a board and can

0272-1732/92/0800-0090$03.00 © 1992 IEEE

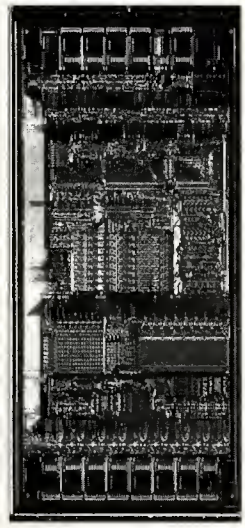
Eight-bit microcontrollers under $1

Motorola's K-series line of controllers for 8-bit microprocessors is priced as low as $1. The chips are aimed at low-cost consumer applications including universal remote controls, keyless automotive entry systems, and auto security systems.

Paul Grimme, product manager for Motorola's CISC (customer-specified integrated circuit) Microcontroller Division, said the series resulted from reducing the feature set of the company's popular 68HC05, to reduce its cost. According to Grimme, the company targeted three groups of customers: users of 4-bit microcontrollers who needed more "horsepower," users of the 05 series who needed a lower cost version, and new customers who could use the inexpensive chip to replace a group of discrete components. The 05 series is itself a lower cost family of 8-bit microprocessors that evolved from the 6800.

The K series includes

- the 68HC05K0, integrating a 32-byte RAM, 504 bytes of ROM, a 15-stage multifunction timer, 10 bidirectional I/O lines (eight with programmable pull-down resistors), a COP (computer operating properly) watchdog timer, low-voltage inhibitor to hold the CPU in reset, power-saving wait and stop modes, and a steering diode on the reset pin;
- the 68HC05K1, combining all the common features of the K0 with 64 bits of EPROM, programmed by user software; and
- the 68HC705K1 incorporating all the features of the K0/K1 with 504 bytes of one-time programmable EPROM that replace the ROM, and an EPROM mask option register for programming all mask options.



Motorola's 68HC05K0 die.

The controllers come in 16-pin PDIP or SOIC packages. *Motorola; $1.00 (K0), $1.50 (K1), $2.50 (705K1) (all in 50,000s).*

Reader Service No. 17

trace and compress data flowing between devices at up to 50 MHz. A scannable 1-Kword × 16-bit RAM stores data. *Texas Instruments; $20 (1,000s).*

Reader Service No. 14

CAD tools

Windows-based PCB design

Max EDS features an "update PCB" command that automatically updates changes from the schematic to the PCB design, adding and deleting parts, connections, and routes. The Windows-based system includes schematic capture, autoplacement, autorouting, and PCB design software. Max EDS operates on Unix and DOS platforms; the manufacturer plans to release an Apple version later this year. *Masstech; $8,750.*

Reader Service No. 15

Integrated IC test solution

The Analytical Probe Station combines a Sun Sparcstation with prober hardware, software, and fixturing to create an environment for testing ICs. The system integrates a floating table with fixturing so the probed device is isolated from mechanical vibration. The station includes a computer-assisted probe, a microscope with motorized mount, and a closed-circuit TV camera and monitor. Sun-based software lets users observe and insert data in an IC's internal nodes while using the CAE/CAD data to guide the probe. *Integrated Measurement Systems; $400,000.*

Reader Service No. 16

Device database on CD-ROM

Designers can search for more than 1.3 million devices and their parameters on the IC/Discrete Parameter Database on CD-ROM. The database includes a software program that automatically matches the user's preferred part numbers against other part numbers in its records. Users can search for a part with either the manufacturer part number or the generic number. The most recent version includes enhanced print features that simplify outputting index information or images. *Information Handling Services; price not given.*

Reader Service No. 18

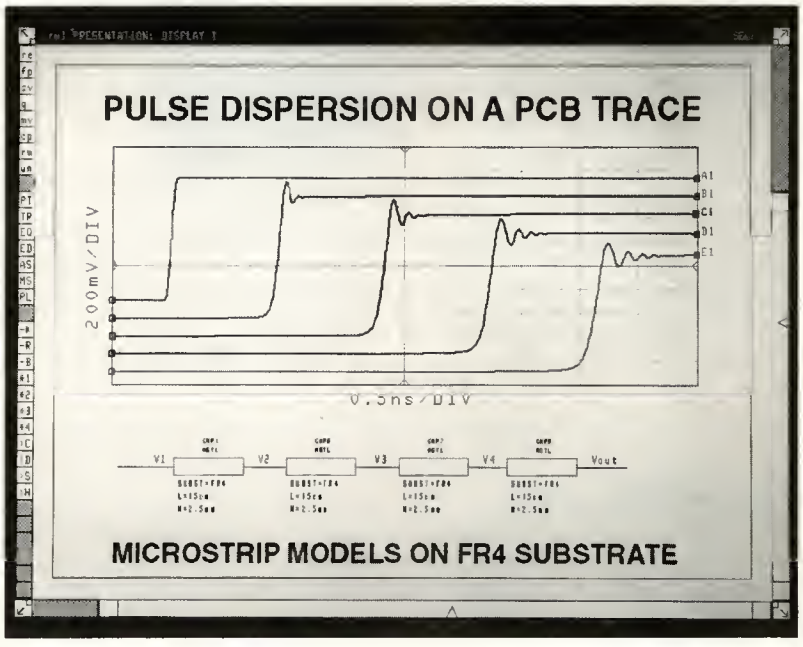
Libraries of 0.6-micron ASIC cells

Two libraries of standard cells and macrocells manufactured in 0.6-micron ASIC technology enhance chip performance by up to 30 percent, according to their manufacturer. The LP600C

High-frequency simulator

Hewlett-Packard says its HP Impulse builds on SPICE's capability to operate with time-domain simulation information by also incorporating frequency-domain information. By doing so, it lets designers access models and measured data previously unavailable, including physical models of transmission media and discontinuities like microstrip, stripline, and coplanar waveguide. Previously, transient circuit simulators modeled equivalent circuit networks to emulate a frequency response. Models using this method are limited because they do not include dispersion or skin-effect loss, which are frequency-dependent effects.

HP Impulse solves problems for high-frequency analog circuit designers working in frequencies from megahertz to terahertz, phase-lock-loop analysis, oscillator startup, and digital-communication circuit analysis. It incorporates frequency-domain components into a time-domain simulator using a technique that the company calls dynamic convolution. This technique converts the frequency response of each frequency domain component into a finite impulse response to obtain the time-domain response. If the circuit has no frequency-domain components, the simulation proceeds in the time domain as it does in a SPICE simulation.



HP Impulse sample output: pulse dispersion on a PCB trace.

HP Impulse is part of HP's Microwave Design System, a software package for RF and microwave engineers who design high-frequency circuits and systems. It is part of a suite of high-frequency design software that includes schematic capture, linear and nonlinear simulation, electromagnetic simulation, active-device modeling, yield analysis, circuit layout and documentation.

The simulator is compatible with most functions of Berkeley SPICE and maintains familiar time-domain response parameters. A SPICE netlist translator assists in the migration of existing designs into the new technology.

Hewlett-Packard; from $31,000.

Reader Service No. 20

CMOS standard-cell library at 3V is designed for portable applications. Chips in the library support speeds up to 70 MHz. If greater speeds are needed (up to 100 MHz), designers can choose from the HP600C library. These 5V devices are aimed at workstations, video equipment, and memory interfaces. Designers can mix 3V and 5V cells within a system.

The design kit contains libraries, application software, and documentation for schematic capture and prelayout solutions. Designers generate netlists and test patterns, then transfer them to the manufacturer for layout. *AT&T Microelectronics; prices vary.*

Reader Service No. 19

Advanced behavioral timing

Release 2.0 software tracks the sophisticated internal timing relationships in complex VLSI devices. It supports timing checks and delays determined by the mode or internal state of operation, since a device's timing characteristics can vary depending on its mode. A complementary product, a 640-pin Device Adapter, expands the manufacturer's line of hardware modelers to anticipate the requirements of complex standard devices, ASICs, and multichip modules. *Logic Modeling; $10,800 (Release 2.0; free upgrades), $9,700 (640-pin Device Adapter).*

Reader Service No. 21

VLSI layout editor

The IC mask design tool VALE (VLSI Advanced Layout Editor) complements IC design environments based on EDA tools from Cadence or Mentor Graphics. VALE's mask design database translates to GDS and CIF formats, enabling two-way interface with other systems. Features include hierarchical design management, editing of subcells in context, and a subprocess interface. VALE supports 32-bit database resolution for VLSI designs with up to 37 layers of mask generation. *Phase Three Logic; $2,995 (end-user license).*

Reader Service No. 22

Signal-processing software and hardware

Signal analysis on Windows or X-windows

Two versions of the DSP Works signal-analysis package support Sun 4 systems and Microsoft Windows. The menu-driven packages generate and process signals and acquire data in real-time.

The Windows version, designed to integrate with the company's QEDesign 1000 and Code Generators, features hardware DSP processor support. The Sun version is available for the Open Look (X-windows) format. Within either version users can create a variety of test signals for detailed analysis including sine, triangular, unit step, square wave, impulse, response, and arbitrary waveform. *Momentum Data Systems; $495 (DSP Works 2.0 for Windows), $2,000 (DSP Works for Sun).*

Reader Service No. 23

Processes real-time audio

With four DSPs on-board and an I/O arrangement that allows one or more daughter boards to access each DSP, the MMI-420 DSP board can process different I/O types simultaneously. Each of the four 27-MHz 56001 DSPs has over 100 Kbytes private memory and can encode 16-bit audio at rates of 48 kHz (digital audio tape) or 44.1 kHz (CDs). Each DSP has a private 16K × 24 SRAM program memory to store downloaded data, a 16K × 24 data memory, and a 16-Kbyte boot EPROM. The MMI-420 occupies a 6U VME slot. *Vigra; $3,225.*

Reader Service No. 24

Test, measurement, and processing on one board

The DT3801 integrated DSP series incorporates test, measurement, and processing systems on a PC AT board based on Texas Instruments' TMS320C40 DSP processor. The boards use all six of the C40's communication ports and the six-channel DMA engine. Data moves through the C40 communication ports, under the supervision of the multichannel DMA controller, while processing occurs. A developer's kit includes TI development tools, Spectron SPOX operating system, and the manufacturer's utility, diagnostic, and demo software. *Data Translation; from $7,195, $2,995 (developer's kit).*

Reader Service No. 25

DSP design system

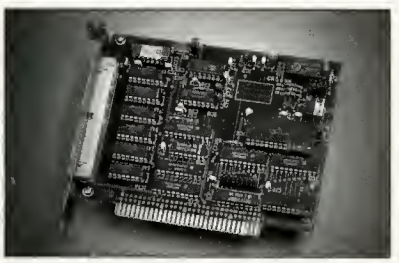
DSP Station, an integrated DSP design system, offers top-down design, from high-level specification through simulation and optimization, into various physical implementations. It features a silicon synthesis capability, including arithmetic optimization and automatic delay minimization. The system provides multiple implementation paths for a designer to choose cost/performance ratios and development and production times. DSP station is available on HP Apollo and Sun Sparc workstations. *Mentor Graphics; $33,000.*

Reader Service No. 26

Analog input card

The AIP-24 analog input card can read up to 24 analog voltages, allowing it to replace several A/D converter boards in many applications. It measures voltages in the ranges of 0 to 10, −5 to +5, or −10 to +10. A 12-bit A/D converter gives a resolution of 1:4000 (0.025 percent). The device processes 16 kHz throughput, yielding conversion times of 25 µs. An on-board programmable gain amplifier gives factors of 100, 10, or unity, allowing users to read millivolt signals accurately. The short-card format occupies a standard I/O slot in DOS machines. *Global Specialties; $495.*

Reader Service No. 27



Global Specialties' AIP-24

Samples up to four channels

Designers who need fast DOS-based instrumentation can use the MAD100 high-speed data acquisition card to simultaneously sample up to four channels. It samples a single channel at up to 100 MHz; in dual- and quad-channel modes it samples at rates up to 50 and 25 MHz. In its "sample accumulate" mode the card takes a programmed number of samples repetitively without having to download the data each time to the PC. Other features of the card include programmable input gain from 640 mV to 64V full scale, memory depth up to 256K, and pretrigger data capability up to the full memory depth. *Markenrich Corp.; $2,995.*

Reader Service No. 28

Reads data up to 225 kHz

The Model 500 DSP board, based on TI's TMS320C51 20-MIPS processor, sustains throughput rates up to 3 Mbytes/s. It reads and writes data at up to a 225-kHz sampling rate. Two 12-bit analog output channels, a buffered digital I/O connector for user expansion, and the two serial interfaces of the TI processor are also on board. Model 500 has 64 Kwords of program RAM and 128 Kwords of dual-ported

Battery charge controller

As computers move off our desks and into our laps and palms, we rely increasingly on rechargeable batteries. Integrated Circuit Systems has developed a controller chip that, the company says, recharges batteries more quickly and avoids the damage caused by overcharging.

The ICS1700 Quick Saver prevents overcharging by automatically switching to a maintenance mode when the battery is fully charged. ICS says this maintenance mode is superior to the conventional trickle charge method with its constant current charge that can encourage growth of crystalline dendrites across the cell plates.

During charging, Quick Saver alternates the charge current with regularly spaced, high-current negative pulses. These pulses prevent damage to the nickel and cadmium plates by breaking up the clusters of oxidized molecules that can form during charging.

Tom Gosse of ICS compares this method of charging to feeding a baby. "Every once in a while, you have to stop and burp the baby, to get out the gas," he says. "Then you can go on feeding."

This reflex system (patented by Christie Electric and licensed to ICS) reduces crystal growth on battery plates that can limit their surface space available for charge, make them resistive to charge, and ultimately cause them to short circuit if the growths reach from the nickel to the cadmium plate.

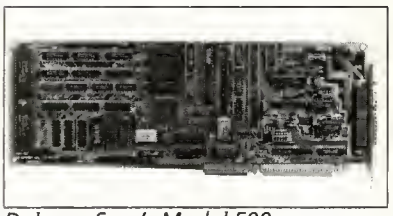
Users can choose charge rate capabilities of 0.5, 1, 2, and 4C. The controller also detects over-temperature through the battery pack thermal switch, short circuits, defective cells, high impedance, and open circuits.

Integrated Circuit Systems; $8.05 (10,000s).

Reader Service No. 32

data RAM. It comes with assembler and debugger software, FFT software, digital filter examples, and a waveform editor. *Dalanco Spry; $1,600.*

Reader Service No. 29



Dalanco Spry's Model 500

Software

Updated TCP/IP program

Version 1.2 of PC/TCP for OS/2 updates the original program with graphical user interfaces, improved error messaging, and enhanced logging of background server activities, according to its developer. The product implements a client-server model with a DOS machine (386 or better) acting as client, server, and workstation.

Exploiting OS/2's multitasking capabilities, it allows machines to access network resources or act as a network resource without interrupting the user's work flow. Enhancements to this version include moving NetBIOS from Ring 3 to Ring 0 and simplifying the installation program. The package includes a version of the Berkeley Internet Naming Daemon that allows users to obtain Internet addresses based on a host name. *FTP Software; $400.*

Reader Service No. 30

Frame-relay-network problems

Two software products help isolate frame-relay-network problems when used with Hewlett-Packard's 4957A, 4957PC, and 4952A wide-area-network protocol analyzers. The HP 18258A frame-relay, data-decode, and statistical analysis software and the HP 18278A frame-relay, post-processing software monitor performance and let users tune network performance. They can track 13 events simultaneously, including the total of frames, frames with bad-frame check sequences, and frames with forward- or backward-congestion bits set. The HP 18258A also tracks and displays any four of the 13 events at runtime and all 13 in post-processing. *Hewlett-Packard; $790 each.*

Reader Service No. 31

Embedded software development tools

A package of microprocessor development tools for the Motorola 680X0, Hitachi H8/300, H8/500, and Gmicro families on the HP Apollo 9000 Series 700 workstations includes ANSI C cross-compilers and Xray Debugger. The compilers comply with the ANSI C standard and accept programs written in C as defined by Kernighan and Ritchie. Xray Debugger debugs optimized C code. *Microtec Research; $4,300 (package includes ANSI C compiler, assembler, linker, and librarian for the 680x0 family).*

Reader Service No. 33

Reader Interest Survey

Indicate your interest in this department by circling the appropriate number on the Reader Service Card.

Low 189 Medium 190 High 191

Micro Law

continued from p. 5

For these reasons, the Court ruled, the ISOs were entitled to go to trial on their claim against tie-ins, under section 1 of the Sherman Act. They were also entitled to go to trial on their claim under section 2 of the Sherman Act, the Court held. Here, Kodak clearly has monopoly power over the Kodak parts market and the Kodak equipment-service market. Kodak's assertion that there cannot be a one-brand market was wrong. Someone who owns a Kodak machine cannot repair it with IBM or Xerox parts, or get IBM or Xerox to repair it.

The Kodak decision has important implications for the electronics and computer industries.

Kodak said that its allegedly monopolistic practices (such as getting OEMs not to sell to ISOs) were justifiable on three grounds:

1) They promoted interbrand competition by allowing Kodak to stress the quality of its service and avoid taking the blame for bad service performed by ISOs.
2) They reduced Kodak's inventory needs.
3) They prevented ISOs from taking a free ride on Kodak's investment.

The Court considered point 1 a mere pretext; the ISOs had evidence to show that customers preferred their service to Kodak's because it was better. Also there was no reason to believe these sophisticated customers were too stupid to know where the blame lay for problems.

The second point was also implausible and did not account for Kodak's pressure on OEMs not to sell to ISOs. Point 3 was unsupported, since Kodak's theory was that service organizations were taking a free ride because they did not enter the equipment and parts businesses as Kodak had. That's not a free ride, the Court said. In fact, forcing new entrants to enter two markets simultaneously creates a barrier to the entry of new competition and is anticompetitive.

Considered as a whole, the Supreme Court's *Kodak* decision seems to be at the opposite pole from the Sixth Circuit's *Virtual Maintenance* decision described earlier. Since the Supreme Court's decisions are the law of the land, the *Kodak* decision has important implications for the electronics and computer industries. Many practices which have developed over the last decade, on the theory that "anything goes" in the current legal and business climate, will now need to be reevaluated for potential antitrust consequences.

Reader Interest Survey

Indicate your interest in this department by circling the appropriate number on the Reader Service Card.

Low 177 Medium 178 High 179

FOR DISPLAY ADVERTISING INFORMATION, CONTACT:

District Manager: D. Rodney Brooks; Tel: (415) 905-0260; Fax: (415) 896-1512.

Recruitment and Classified Advertising: D. Rodney Brooks; Tel: (415) 905-0260; Fax (415) 896-1512.

Director of Sales: Randall L. Stickrod, 544 Second St., Suite 200, San Francisco, CA 94107; Tel: (415) 905-0260; Fax: (415) 896-1512.

For production information, conference, and classified advertising, contact Marian Tibayan.

IEEE MICRO, 10662 Los Vaqueros Cir., PO Box 3014, Los Alamitos, CA 90720-1264; phone (714) 821-8380; fax (714) 821-4010.

THE FOLLOWING INFORMATION IS AVAILABLE:

Contact the Publications Office;
to facilitate handling, please request by number.

- Membership application, student #203, others #202
- Periodicals subscription form for individuals #200
- Periodicals subscription form for organizations #199
- Publications catalog #201
- Compmail electronic mail brochure #194
- Technical committee list/application #197
- Chapters lists, start-up procedures #193
- Student scholarship information #192
- Volunteer leaders/staff directory #196
- IEEE senior member grade application #204
  (requires ten years practice and significant performance in five of those ten)

To check membership status or report a change of address, call the IEEE toll-free number, 1-800-678-4333. Direct all other Computer Society-related questions to the Publications Office.

PURPOSE

The IEEE Computer Society advances the theory and practice of computer science and engineering, promotes the exchange of technical information among 100,000 members worldwide, and provides a wide range of services to members and nonmembers.

MEMBERSHIP

Members receive the acclaimed monthly magazine *Computer*, discounts, and opportunities to serve (all activities are led by volunteer members). Membership is open to all IEEE members, affiliate society members, and others interested in the computer field.

IEEE COMPUTER SOCIETY®

A member society of the
Institute of Electrical and Electronics Engineers, Inc.

PUBLICATIONS AND ACTIVITIES

Computer. An authoritative, easy-to-read magazine containing tutorial and in-depth articles on topics across the computer field, plus news, conferences, calendar, interviews, and product reviews.

Periodicals. The society publishes seven magazines and five research transactions. Refer to membership application or request information as noted at left.

Conference Proceedings, Tutorial Texts, Standards Documents. The Computer Society Press publishes more than 100 titles every year.

Standards Working Groups. More than 100 of these groups produce IEEE standards used throughout the industrial world.

Technical Committees. More than 30 TCs publish newsletters, provide interaction with peers in specialty areas, and directly influence standards, conferences, and education.

Conferences/Education. The society holds about 100 conferences each year and sponsors many educational activities, including computing science accreditation.

Chapters. Regular and student chapters worldwide provide the opportunity to interact with colleagues, hear technical experts, and serve the local professional community.

OMBUDSMAN

Members experiencing problems — magazine delivery, membership status, or unresolved complaints — may write to the ombudsman at the Publications Office.

EXECUTIVE COMMITTEE

President: 8ruce D. Shriver*
17 8ethea Drive
Ossining, NY 10562
Phone: (914) 762-3251
Fax: (914) 941-9181

President-Elect: James H. Aylor*
Past President: Duncan H. Lawrie*

VP, Conferences and Tutorials: 8arry W. Johnson (1st VP)*
VP, Educational Activities: Raymond E. Miller (2nd VP)*
VP, Membership Activities: Florenza C. Albert-Howard*
VP, Press Activities: Yale N. Patt*
VP, Publications: Harold S. Stone*
VP, Standards Activities: Gary Robinson†
VP, Technical Activities: Joseph 8oykin†

Secretary: Ronald G. Hoelzeman*
Treasurer: Laurel V. Kaleda†
IEEE Division V Director: 8ill D. Carroll†
IEEE Division VIII Director: Helen M. Wood†
Executive Director: T. Michael Elliott†

*voting member of the Board of Governors
†nonvoting member of the Board of Governors

BOARD OF GOVERNORS

Term Expiring 1992:
Alicja I. Ellis, Ronald G. Hoelzeman, Tadao Ichikawa, C.V. Ramamoorthy, Sallie V. Sheppard, Harold Stone, Akihiko Yamada

Term Expiring 1993:
Florenza Albert-Howard, Jon T. 8utler, Michael C. Mulder, Yale N. Patt, Anneliese von Mayrhauser, Benjamin W. Wah, Ronald Waxman

Term Expiring 1994:
Mario R. 8arbacci, Luis-Felipe Cabrera, Wolfgang K. Giloi, Guylaine M. Pollock, John P. Riganati, Ronald D. Williams, Thomas W. Williams

Next Board Meeting
November 20, 1992, 8:30 a.m.
Minneapolis Marriott, Minneapolis, Minn.

SENIOR STAFF

Executive Director: T. Michael Elliott
Publisher: H. True Seaborn
Director, Conferences and Tutorials: Anne Marie Kelly
Director, Finance and Information Services: Debroah Rafal
Director, 8oard and Administrative Services: Violet S. Doan
Assistant to the Executive Director: Sandra K. Pfau

COMPUTER SOCIETY OFFICES

Headquarters Office
1730 Massachusetts Ave. NW
Washington, DC 20036-1903
Phone: (202) 371-0101
Fax: (202) 728-9614

Publications Office
10662 Los Vaqueros Cir.
PO Box 3014
Los Alamitos, CA 90720-1264
Membership and General Information:
(714) 821-8380
Publication Orders: (800) 272-6657
Fax: (714) 821-4010

European Office
13, Ave. de L'Aquilon
8-1200 8russels, 8elgium
Phone: 32 (2) 770-21-98
Fax: 32 (2) 770-85-05

Asia/Pacific Office
Doshima Building
2-19-1 Minami-Aoyama, Minato-ku
Tokyo 107, Japan
Phone: 81 (3) 3408-3118
Fax: 81 (3) 3408-3553

IEEE OFFICERS

President: Merrill W. 8uckley, Jr.
President-Elect: Martha Sloan
Past President: Eric E. Sumner
Secretary: Karsten E. Drangeid
Treasurer: Theodore W. Hissey, Jr

VP, Educational Activities: Edward A. Parrish
VP, Professional Activities: Arvid G. Larson
VP, Publication Activities: James T. Cain
VP, Regional Activities: Luis T. Gandia
VP, Standards Activities: Marco W. Migliaro
VP, Technical Activities: Fernando Aldana

NEW BOOKS

from IEEE Computer Society Press

Groupware:
Software for Computer-Supported Cooperative Work

edited by David Marca and Geoffrey Bock

This book focuses on the task of bringing computing face-to-face with groups of people as they work. It concentrates on the designing of software to fit the way groups interact in specific work situations. The tutorial covers the social and technical aspects of groupware development and presents a wide range of material on the need to design group-related computer and information systems.

Sections Include: Groups and Groupware, Conceptual Frameworks, Design Methods, System-Related and UI-Related Enabling Technologies, Computer Supported Meetings, Bridging Time and Space, Coordinators.

592 pages. August 1992. Hardcover. ISBN 0-8186-2637-2.
Catalog # 2637 — $75.00 Members $45.00

Real-Time Systems Design and Analysis

by Phillip A. Laplante

This monograph provides a comprehensive guide to the practical design and analysis of real-time systems. The book covers all aspects of real-time software design including computer architecture, operating systems, programming languages, software engineering, and systems integration. It contains many examples, illustrations, and exercises as well as practical tools which the software engineer or student can apply today in the field or laboratory.

Sections Include: Computer Hardware Technologies; Language Issues; Real-Time Modeling Techniques; Real-Time Kernels; Inter-Task Communication and Synchronization; Real-Time Memory Management; System Performance Analysis and Optimization; Testing, Fault Tolerance, and Reliability; Multiprocessing Systems; Hardware/Software Integration.

c.350 pages. September 1992. Hardcover. ISBN 0-7803-0402-0.
*Catalog # 3107 — * $39.95 Members $32.00*
(* prepublication price)

Knowledge-Based Systems:
Fundamentals and Tools

edited by Oscar N. Garcia and Yi-Tzuu Chien

This tutorial emphasizes the characteristics of knowledge-based systems, including rapid prototyping, knowledge engineering, representation paradigms, logical foundations, inferencing modes, verification and validation of knowledge bases, and languages, shells, and tools. Its 35 papers explore recent research and provide a brief account of the status of today's knowledge-based systems. The book considers how to match the appropriate method to an existing problem; covers eight paradigms used in practice today; and introduces the terminology of logic-based database development.

Sections Include: The Role of Knowledge Engineering; Logic, Logic Programming, and Prolog; Rule-Based Systems — Forward and Backward Chaining; Dealing with Uncertainty; Learning and the Connectionist Approach; Knowledge Validation, Verification, and Utilization.

512 pages. January 1992. Softcover. ISBN 0-8186-1924-4.
Catalog # 1924 — $65.00 Members $45.00

Information Systems and Decision Processes

by Edward A. Stohr and Benn R. Konsynski

This book contains the collaborative studies of DSS researchers and explains the important opportunities and problems in the application of information systems to the decision process in organizations. It also describes the multidisciplinary approaches developed by these research groups and synthesizes the ideas from each field to allow for a fuller understanding of information systems in organizations. The summaries and conclusions from their studies are contained in this monograph.

Sections: Review and Critique of DSS, Decision Processes, Behavioral Decision Theory and DSS, Group Decision Support Systems, Organizational Decision Support Systems, Technology Environments to Support Decision Processes, Model Management Systems, Research Challenges, Research Approaches in ISDP.

368 pages. August 1992. Hardcover. ISBN 0-8186-2802-2.
Catalog # 2802 — $65.00 Members $35.00

Call toll-free to order: 1-800-CS-BOOKS
or order by *FAX*: (714) 821-4010
(in CA call (714) 821-8380)